



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 98584

TO: Mark Navarro
Location: 8A15
Art Unit: 1645
Friday, July 11, 2003

Case Serial Number: 10/091442

From: Barb O'Bryen
Location: Biotech-Chem Library
CM1-6A05
Phone: 308-4291 *BOB*

barbara.obryen@uspto.gov

Search Notes

O'Bryen, Barbara

From: Navarro, Albert
Sent: Thursday, July 10, 2003 12:47 PM
To: O'Bryen, Barbara
Subject: 10/091,442

Mark Navarro
1645
306-3225
8A15

Hello Barb,

Could I trouble you for a search and an interference search of:

SEQ ID NO: 1-9 and 34

Thanks

Mark

7/14/93
(7/17/92 FP)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 8.29522 Seconds
(without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-3

Perfect score: 100

Sequence: 1 VERVDFTNHLEDFRRNINK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP nhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	71.0	380	11	Q9D695	Q9D695 mus musculus
2	68.0	380	11	Q920J5	Q920J5 rattus norv
3	51.5	509	12	Q9YQ03	Q9YQ03 myxoma viru
4	51.5	509	12	Q9PX36	Q9PX36 myxoma viru
5	50.0	338	4	Q9BYR8	Q9BYR8 homo sapien
6	50.0	2020	5	Q8T2M1	Q8T2M1 dictyosteli
7	48.0	380	5	Q8WQX0	Q8WQX0 rhinipicephal
8	48.0	472	5	Q09657	Q09657 caenorhabdi
9	47.0	245	16	Q8XM42	Q8XM42 clostridium
10	47.0	419	4	Q43611	Q43611 homo sapien
11	47.0	672	5	Q25586	Q25586 onchocerca
12	47.0	1024	4	Q14527	Q14527 homo sapien
13	47.0	1046	4	Q9UD76	Q9UD76 homo sapien
14	47.0	1046	11	Q920H8	Q920H8 mus musculu
15	47.0	1046	11	Q55156	Q55156 rattus norv
16	47.0	1047	11	Q9EP81	Q9EP81 mus musculu

17	46.0	173	10	Q946D0	Q946D0 theobroma c
18	46.0	174	10	Q945Y0	Q945Y0 theobroma c
19	46.0	231	4	Q8WM89	Q8WM89 homo sapien
20	46.0	369	4	Q9BYF7	Q9BYF7 homo sapien
21	46.0	427	5	Q9U114	Q9U114 drosophila
22	46.0	427	5	Q9VFC2	Q9VFC2 drosophila
23	45.0	309	16	Q92PT5	Q92PT5 rhizobium m
24	45.0	361	4	Q9BX29	Q9BX29 homo sapien
25	45.0	443	5	Q76878	Q76878 drosophila
26	45.0	539	11	Q9DCD5	Q9DCD5 mus musculu
27	45.0	673	16	Q24911	Q24911 helicobacte
28	45.0	673	16	Q9ZMW7	Q9ZMW7 helicobacte
29	44.5	330	16	Q97FL1	Q97FL1 clostridium
30	44.0	311	2	Q9WTC2	Q9WTC2 escherichia
31	44.0	348	16	Q9PR45	Q9PR45 ureaplasma
32	44.0	369	10	Q9LY90	Q9LY90 arabidopsis
33	44.0	542	10	Q9FQY7	Q9FQY7 capsicum an
34	44.0	617	2	Q32544	Q32544 escherichia
35	44.0	709	5	Q9GRG1	Q9GRG1 tetrahymena
36	43.5	119	16	Q55617	Q55617 synecocyst
37	43.0	58	16	Q97Q04	Q97Q04 streptococ
38	43.0	173	11	Q08882	Q08882 rattus norv
39	43.0	175	5	Q9VXT8	Q9VXT8 drosophila
40	43.0	259	16	Q9RQW0	Q9RQW0 neisseria m
41	43.0	301	5	Q95PD8	Q95PD8 calliphora
42	43.0	301	16	Q25565	Q25565 helicobacte
43	43.0	363	16	Q9ZKJ2	Q9ZKJ2 helicobacte
44	43.0	388	16	Q8XN50	Q8XN50 clostridium
45	43.0	468	10	Q9S1B7	Q9S1B7 arabidopsis

ALIGNMENTS

Q9D695	09D695	PRELIMINARY;	PRT;	380 AA.
ID	Q9D695	01-JUN-2001 (TRENBLREL. 17, Created)		
AC	Q9D695	01-JUN-2001 (TRENBLREL. 17, Last sequence update)		
DT	01-JUN-2001 (TRENBLREL. 17, Last sequence update)			
DT	01-JUN-2002 (TRENBLREL. 21, Last annotation update)			
DE	4631416M05R1K protein (Megsin).			
GN	SERPINB7 OR 4631416M05R1K.			
OS	Mus musculus (Mouse).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=SKIN;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,			
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo T., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyon P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Winking L.,			
RA	Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			

```

RX MEDLINE-21368006; Pubmed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Kurokawa K.;
RA Yagi M., Naganu N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
DR EMBL: AK014524; BAB29410.1; -.
DR HSSP: P05121; 1A7C.
DR MGD: MGI:2151053; Serpinb7.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; Serpin_1.
DR SMART: SM00093; SERPIN_1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCF9CF4 CRC64;

Query Match          71.0%; Score 71; DB 11; Length 380;
Best Local Similarity 73.7%; Pred. No. 0.0027;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDTNRNIN 19
DB 126 VERVDFTNDYQDTRFKIN 144

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
AC 0920J5;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Megsin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21368006; Pubmed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Naganu N., Inagi R., Kurokawa K.;
RT "Cloning of rodent meglin revealed its up-regulation in
RT mesangio proliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AF105329; AAL16769.1; -.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; Serpin_1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 380 AA; 42821 MW; D8076CABEE2C2FBC CRC64;

Query Match          68.0%; Score 68; DB 11; Length 380;
Best Local Similarity 68.4%; Pred. No. 0.0081;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDTNRNIN 19
DB 126 VERVDFTNDYQDTRFKIN 144

RESULT 3
Q9Y003 PRELIMINARY; PRT; 509 AA.
AC 09Y003;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE M-T6 protein.
GN MT6.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

```

```

OC Leporipoxvirus.
OX NCBI_TaxID=31530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LAUSANNE;
RA Salim S., Meissner K., Neubert A., Doehner L., Becher D.;
RT "DNA sequence of the M-T6 gene of Myxoma virus strain Lausanne.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ012283; GAN09975.1; -.
DR InterPro: IPR000210; BTR_P0Z.
DR InterPro: IPR000875; Cectropin.
DR InterPro: IPR001798; Kelch.
DR Pfam: PF00651; BTR; 1.
DR Pfam: PF01344; Kelch; 5.
DR SMART: SM00225; BTR; 1.
DR PROSITE: PS50097; BTR; 1.
DR PROSITE: PS00268; CECTROPIN; UNKNOWN_1.
SQ SEQUENCE 509 AA; 58166 MW; 5A7799B6015C6BFA CRC64;

Query Match          51.5%; Score 51.5; DB 12; Length 509;
Best Local Similarity 41.4%; Pred. No. 5.2;
Matches 12; Conservative 3; Mismatches 3; Indels 11; Gaps 1;

QY 1 VERVDFTN-----HLEDTRNIN 18
DB 109 VSRVDFTNCVSARFQFAETVHIEDLKRLN 137

RESULT 4
Q9PX36 PRELIMINARY; PRT; 509 AA.
AC 09PX36;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE M006L protein.
GN M006L OR M006R.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
OX NCBI_TaxID=31530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LAUSANNE;
RX MEDLINE-20032073; Pubmed-10562494;
RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,
RA Macaulay C., Miller D., Evans D., McFadden G.;
RT "The complete DNA sequence of myxoma virus.";
RL Virology 264:298-318(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LAUSANNE;
RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,
RA Macaulay C., Miller D., Evans D., McFadden G.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF170726; AAF15052.1; -.
DR EMBL: AF170726; AAF14893.1; -.
DR InterPro: IPR000210; BTR_P0Z.
DR InterPro: IPR000875; Cectropin.
DR InterPro: IPR001798; Kelch.
DR Pfam: PF00651; BTR; 1.
DR Pfam: PF01344; Kelch; 5.
DR SMART: SM00225; BTR; 1.
DR PROSITE: PS50097; BTR; 1.
DR PROSITE: PS00268; CECTROPIN; UNKNOWN_1.
SQ SEQUENCE 509 AA; 58166 MW; 5D0129AA03DE4B7D CRC64;

Query Match          51.5%; Score 51.5; DB 12; Length 509;
Best Local Similarity 41.4%; Pred. No. 5.2;
Matches 12; Conservative 3; Mismatches 3; Indels 11; Gaps 1;

QY 1 VERVDFTN-----HLEDTRNIN 18
DB 109 VSRVDFTNCVSARFQFAETVHIEDLKRLN 137

```

DB 109 VSRVDPNTCVSAFOFAETVHIEDIKRNLN 137

RESULT 5

Q9BYF8 PRELIMINARY; PRT; 338 AA.
AC Q9BYF8; 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
SCCAB.
GN SCCAB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Suminami Y., Kishi F., Murakami A., Sakaguchi Y., Kato H.;
RT "Novel Forms of SCC Antigen Transcripts Produced by Alternative
RT Splicing";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR HSSP; P01008; IATH.
DR EMBL; AB046399; BAB40772.1; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 338 AA; 38519 MW; 5AB9CEFA80C1ABDF CRC64;

Query Match 50.0%; Score 50; DB 4; Length 338;
Best Local Similarity 55.6%; Pred. No. 5.8;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VERVDFTNHLEDTNRNIN 18

DB 131 VESYDFANAPESKKNIN 148

RESULT 6

ID 08T2M1 PRELIMINARY; PRT; 2020 AA.
AC 08T2M1; 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Hypothetical 232.6 kDa protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC Gloeckner G., Eichinger L., Szafranski K., Pachepat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Gilgo R., Kumpf K.,
RA Tunngal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostellum";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115594; AAL92302.1; -
KW Hypothetical protein.
SQ SEQUENCE 2020 AA; 232615 MW; 01E221CE63114DD3 CRC64;

Query Match 50.0%; Score 50; DB 5; Length 2020;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 VDFTNHLEDTNRNIN 19

DB 967 IDDSNHSDDTJNSINK 982

RESULT 7

Q8WQX0 PRELIMINARY; PRT; 380 AA.
AC Q8WQX0; 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
OS Serine proteinase inhibitor serpin-2.
OC Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
OX NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A.
RA Mielenz A., Sugimoco C., Onuma M.;
RT "Blood meal feeding-induced serine proteinase inhibitor-2 from
RT Rhipicephalus appendiculatus";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035780; AAK61376.1; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN; 1.
SQ SEQUENCE 380 AA; 42729 MW; 762B7CA4E1F611F2 CRC64;

Query Match 48.0%; Score 48; DB 5; Length 380;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 VERVDFTNHLEDTNRNIN 18

DB 120 IESYDFKTHYKVRQGVN 137

RESULT 8

ID 009657 PRELIMINARY; PRT; 472 AA.
AC 009657; 01-JAN-1999 (TReMBLrel. 09, Created)
DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Hypothetical 51.4 kDa protein ZK1320.9 in chromosome II.
GN ZK1320.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Berks M.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO C.KLUYVERI ORFZ.
DR EMBL; Z46934; CAA87047.1; -
DR WormPep: ZK1320.9; CE01707.
DR InterPro: IPR003702; AcetylCoA_hydro.
DR InterPro: IPR001220; Lectin_legh.
DR Pfam: PF02550; AcetylCoA_hydro.1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN.1.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51369 MW; 08C3351E1B7C348 CRC64;

Query Match 48.0%; Score 48; DB 5; Length 472;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 DFTNHLEDTNRN 16

DB 328 DFTNHLDIVARN 339

RESULT 9

Q8XM42

ID 08XM42 PRELIMINARY; PRT; 245 AA.
 AC 08XM42;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein CPE0848.
 GN CPE0848.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL; AP003188; BAB80554.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 245 AA; 28398 MW; 229703EAC24770D2 CRC64;

Query Match 47.0%; Score 47; DB 16; Length 245;
 Best Local Similarity 57.1%; Pred. No. 12;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 5 DFTNHLDETRRNIN 18
 Db 113 DLSNLSIENRRNIN 126

RESULT 10

ID 043611 PRELIMINARY; PRT; 419 AA.

AC 043611;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Probable CLIP protein (Fragment).
 GN WSCR4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96411691; PubMed=8812460;
 RA Osborne L.R., Martindale D.W., Scherer S.W., Shi X.-M., Hutzenga J.,
 RA Heng H.H.Q., Costa T., Pober B., Lew L., Brinkman J., Rommens J.,
 RA Koop B.F., Tsui L.-C.;
 RT "Identification of genes from a 500-kb region at 7q11.23 that is
 RT commonly deleted in Williams syndrome patients.";
 RL Genomics 36:328-336(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Martindale D.W., Koop B.F.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF041055; AAB96784.1; JOINED.
 DR EMBL; AF041056; AAB96784.1; JOINED.
 DR EMBL; AF041057; AAB96784.1; JOINED.
 DR EMBL; AF041058; AAB96784.1; JOINED.
 DR InterPro: IPR000938; CAP-Gly.
 DR Pfam: PF01302; CAP_GLY; 2.
 DR PROSITE: PS00845; CAP_GLY_1; UNKNOWN_2.
 FT NON_TER 1
 FT SEQUENCE 419 AA; 44825 MW; 547CA46012255C25 CRC64;

Query Match 47.0%; Score 47; DB 4; Length 419;
 Best Local Similarity 50.0%; Pred. No. 22;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 OY 2 ERVDFTNHLDETRRNIN 17
 Db 386 EKVDSLNSQLEERKRV 401

RESULT 11

ID 025586 PRELIMINARY; PRT; 672 AA.

AC 025586;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Paramyosin-related protein (Fragment).
 OS Onchocerca gibsoni.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6284;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96360163; PubMed=8719949;
 RA Zhang D., Miller D.J.;
 RT "Characterization of a novel non-muscle myosin-related protein from
 RT Onchocerca gibsoni.";
 RL Int. J. Parasitol. 25:1385-1391(1995).
 DR EMBL; U20609; AAB50272.1;
 DR InterPro: IPR000727; T-SNARE.
 FT NON_TER 1
 FT SEQUENCE 672 AA; 80074 MW; C610249CEFFFD1 CRC64;

Query Match 47.0%; Score 47; DB 5; Length 672;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRR 15
 Db 83 ISVDVDFANQLEDERR 97

RESULT 12

ID 014527 PRELIMINARY; PRT; 1024 AA.

AC 014527;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE KIAA0291 protein (Fragment).
 GN KIAA0291.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Ohara O., Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N.,
 RA Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB006629; BAA22960.2;
 DR InterPro: IPR000938; CAP-Gly.
 DR Pfam: PF01302; CAP_GLY; 2.
 DR PROSITE: PS00845; CAP_GLY_1; UNKNOWN_2.
 FT NON_TER 1
 FT SEQUENCE 1024 AA; 112981 MW; 12278FC36788B620 CRC64;

Query Match 47.0%; Score 47; DB 4; Length 1024;
 Best Local Similarity 50.0%; Pred. No. 60;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 OY 2 ERVDFTNHLDETRRNIN 17

Db 440 EKVDLSNLEEEERRKV 455

RESULT 13

ID 09UDT6 PRELIMINARY; PRT: 1046 AA.
 G9UDT6
 AC 09UDT6
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE KIA0291, similar rodent cytoplasmic linker protein CLIP-115 and
 reslin.
 GN WUSC:H.GS166C05.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99063792; PubMed-9847074;
 RT "Toward a complete human genome sequence."
 RL Genome Res. 8:1097-1108(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bemis G., Langston Y., Kellen J.;
 RT "The sequence of Homo sapiens BAC clone GS1-166C5."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005015; AAF03524.1;
 DR InterPro: IPR000938; CAP_Gly; 2.
 DR Pfam: PF01302; CAP_Gly; 2.
 DR PROSITE: PS00845; CAP_Gly_1; UNKNOWN_2.
 SQ SEQUENCE 1046 AA; 115837 MW; 7A0B3C796E1C6E25 CRC64;
 Query Match 47.0%; Score 47; DB 4; Length 1046;
 Best Local Similarity 50.0%; Pred. No. 61;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 ERVDFTNHLDPTRRNI 17
 Db 427 EKVDLSNLEEEERRKV 442
 RESULT 14
 G9ZOH8 PRELIMINARY; PRT: 1046 AA.
 AC 09ZOH8
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE CLIP-115.
 GN CYLN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE-99017971; PubMed-9799601;
 RA Hoogenraad C.C., Eussen B.H.J., Langeveld A., van Haperen R.,
 RA Witterberg S., Mouters C.H., Grosveld F., de Zeeuw C.I., Galjart N.;
 RT "The murine CYLN2 gene: genomic organization, chromosome localization,
 RT and comparison to the human gene that is located within the 7q11.23
 RT Williams syndrome critical region."
 RL Genomics 53:348-358(1998).
 DR EMBL; AJ228865; CAA13069.1;
 DR EMBL; AJ228866; CAA13069.1; JOINED.

DR EMBL; AJ228867; CAA13069.1; JOINED.
 DR EMBL; AJ228868; CAA13069.1; JOINED.
 DR EMBL; AJ228869; CAA13069.1; JOINED.
 DR EMBL; AJ228870; CAA13069.1; JOINED.
 DR EMBL; AJ228871; CAA13069.1; JOINED.
 DR EMBL; AJ228872; CAA13069.1; JOINED.
 DR EMBL; AJ228873; CAA13069.1; JOINED.
 DR EMBL; AJ228874; CAA13069.1; JOINED.
 DR EMBL; AJ228875; CAA13069.1; JOINED.
 DR EMBL; AJ228876; CAA13069.1; JOINED.
 DR EMBL; AJ228877; CAA13069.1; JOINED.
 DR EMBL; AJ228878; CAA13069.1; JOINED.
 DR EMBL; AJ228879; CAA13069.1; JOINED.
 DR EMBL; AJ228880; CAA13069.1; JOINED.
 DR EMBL; AJ228883; CAA13069.1; JOINED.
 DR MGD; MGI:131316; Cyln2.
 DR InterPro: IPR000938; CAP_Gly.
 DR InterPro: IPR000812; TFIIB_euk.
 DR Pfam: PF01302; CAP_Gly; 2.
 DR PROSITE: PS00845; CAP_Gly_1; UNKNOWN_1.
 DR PROSITE: PS00782; TFIIB; UNKNOWN_1.
 SQ SEQUENCE 1046 AA; 115494 MW; A2B9C28FC95B808C CRC64;
 Query Match 47.0%; Score 47; DB 11; Length 1046;
 Best Local Similarity 50.0%; Pred. No. 61;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 ERVDFTNHLDPTRRNI 17
 Db 428 EKVDLSNLEEEERRKV 443

RESULT 15
 ID 055156 PRELIMINARY; PRT: 1046 AA.
 AC 055156
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE CLIP-115 protein.
 GN CLIP-115.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98087115; PubMed-9427243;
 RA de Zeeuw C.I., Hoogenraad C.C., Goedknegt E., Hertzberg E.,
 RA Neubauer A., Grosveld F.G., Galjart N.J.;
 RT "CLIP-115, a novel brain specific cytoplasmic linker protein, mediates
 RT the localisation of dendritic lamellar bodies."
 RL Neuron 19:1187-1199(1997).
 DR EMBL; AJ000485; CAA04123.1;
 DR InterPro: IPR000938; CAP_Gly.
 DR InterPro: IPR000812; TFIIB_euk.
 DR Pfam: PF01302; CAP_Gly; 2.
 DR PROSITE: PS00845; CAP_Gly_1; UNKNOWN_2.
 DR PROSITE: PS00782; TFIIB; UNKNOWN_1.
 SQ SEQUENCE 1046 AA; 115480 MW; 72E6CE9F76D2A1D0 CRC64;
 Query Match 47.0%; Score 47; DB 11; Length 1046;
 Best Local Similarity 50.0%; Pred. No. 61;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 2 ERVDFTNHLDPTRRNI 17
 Db 428 EKVDLSNLEEEERRKV 443

Search completed: July 11, 2003, 11:58:54
 Job time : 11.2952 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 9.75468 seconds
(without alignments)
232.223 Million cell updates/sec

Title: US-10-091-442-4

Perfect score: 83

Sequence: 1 SYIEVTEEGTEATATG 17

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	380	15	AA18379 Human megakaryocyte
2	83	100.0	380	20	AA108254 Human megaslin prote
3	83	100.0	380	21	AA124142 Human megaslin prote
4	83	100.0	380	22	AA183075 Human megaslin prote
5	69	83.1	17	15	AA183110 Human megakaryocyte
6	69	83.1	217	20	AA129329 Human secreted pro
7	69	83.1	217	22	AA1039053 Human secreted pro
8	69	83.1	217	23	AA185764 Human polypeptide
9	69	83.1	340	22	AA1016928 Human novel secret
10	69	83.1	368	20	AA108256 Mouse megaslin prote

11	69	83.1	368	21	AA124151
12	69	83.1	368	22	AA183077
13	69	83.1	379	21	AA151515
14	69	83.1	617	21	AA151516
15	67	80.7	135	22	AA172661
16	67	80.7	316	20	AA128644
17	67	80.7	391	20	AA141029
18	67	80.7	391	20	AA125926
19	67	80.7	391	21	AA111314
20	67	80.7	391	21	AA110953
21	67	80.7	391	21	AA110954
22	67	80.7	391	22	AA172651
23	67	80.7	391	23	AA172651
24	67	80.7	400	20	AA141030
25	67	80.7	400	21	AA111315
26	67	80.7	400	22	AA172652
27	67	80.7	400	23	AA172652
28	67	80.7	1120	22	AA1828279
29	66	79.5	392	22	AA1828279
30	66	79.5	424	22	AA1828279
31	66	79.5	431	22	AA1828279
32	65	78.3	380	20	AA108255
33	65	78.3	380	21	AA124150
34	65	78.3	380	22	AA124150
35	65	78.3	380	22	AA124150
36	62	74.7	116	22	AA102921
37	62	74.7	188	22	AA102921
38	62	74.7	372	22	AA102921
39	62	74.7	374	22	AA102921
40	62	74.7	376	22	AA102921
41	62	74.7	376	22	AA102921
42	62	74.7	379	16	AA164159
43	62	74.7	379	17	AA164159
44	62	74.7	389	22	AA102973
45	62	74.7	400	22	AA102962

ALIGNMENTS

RESULT 1	AA18379	standard; Protein: 380 AA.
AA18379:		
AC	AA18379:	
DT	16-AUG-1994	(first entry)
XX		
DE	Human megakaryocyte differentiation factor.	
XX		
KW	Human megakaryocyte differentiation factor; MGF; thrombopoietin;	
KW	haematopoietic stimulating factor; thrombocytopoietin; platelet;	
KW	bone marrow transplantation; cancer chemotherapy.	
XX		
OS	Homo sapiens.	
XX		
PN	EP583864-A.	
XX		
PD	23-FEB-1994.	
XX		
PF	19-JUL-1993;	93EP-0305654.
XX		
PR	17-JUL-1992;	92JP-0212305.
PR	04-MAR-1993;	93JP-0067339.
XX		
PA	(SUNR) SUNTORY LTD.	
XX	(TSUJ) TSUJIMOTO M.	
PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;	
PI	Tsujimoto M, Tsunokawa N, Yamaguchi N, Yamachi K;	
XX		
DR	WPI: 1994-058782/08.	
DR	N-PSDB; AA056670.	

```

XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
PS Claim 7; Page 30-32; 47pp; English.
XX
CC Human MDF can be isolated from a culture of human epidermoid
CC carcinoma A431 cells in protein-free medium. The MDF stimulates
CC differentiation of megakaryocytes from myeloid cells in the presence
CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC for treatment of diseases involving a decrease in platelet number
CC (esp. thrombocytopenia) such as occurs in bone marrow
CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by
CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 83; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEECTEATATG 17
DB 325 SYIEVTEECTEATATG 341

RESULT 2
AA08254
ID AAY08254 standard; Protein; 380 AA.
XX
AC AAY08254;
XX
DF 14-JUL-1999 (first entry)
XX
DE Human megsin protein.
XX
KW Megsin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
KW human; rat; murine.
XX
OS Homo sapiens.
XX
PN WO9915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KUROO) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 1999-276983/23.
DR N-PSDB: AAX56712.
XX
XX Megsin protein expressed specifically in mesangial cells
PS Claim 1; Page 62-64; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megsin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as Iga nephropathy.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 83; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 SYIEVTEECTEATATG 17
DB 325 SYIEVTEECTEATATG 341

RESULT 3
AAB24142
ID AAB24142 standard; Protein; 380 AA.
XX
AC AAB24142;
XX
DF 30-JAN-2001 (first entry)
XX
DE Human megsin protein sequence SEQ ID NO:2.
XX
KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
KW Iga; Immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
OS Homo sapiens.
XX
PN WO200057189-A1.
XX
PD 28-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-JP01646.
XX
PR 19-MAR-1999; 98JP-0075305.
PR 28-OCT-1999; 99JP-0306623.
XX
PA (KUROO) KUROKAWA K.
PA (FUSO) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 2000-611642/58.
DR N-PSDB: AAA99294.
XX
XX Evaluating renal function comprises assaying megsin protein in
PT biological sample
XX
PS Example 2; Page 66-69; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megsin protein in biological sample. Also
CC described are: (1) use of a anti-megsin protein antibody for diagnosing
CC renal function; and (2) a kit for detecting megsin protein comprising:
CC (a) anti-megsin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying megsin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence represents the human megsin
CC protein, which is given in the exemplification of the present invention.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 83; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEECTEATATG 17
DB 325 SYIEVTEECTEATATG 341

RESULT 4
AAB83075
ID AAB83075 standard; Protein; 380 AA.
XX
AC AAB83075;

```



```

XX 10-JUL-2001 (first entry)
XX Human megsln protein.
DE
XX Human: megsln; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening.
XX
OS Homo sapiens.
XX
PN WO200124628-A1.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000MO-JP06988.
XX
PR 06-OCT-1999; 99JP-0285736.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T.
XX
DR WPI: 2001-300136/31.
DR N-PSDB; AAF82438.
XX
PT Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments -
XX
PS Example 4; Page 44-46; 62pp; Japanese.
XX
CC The present sequence is human megsln. The human megsln coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.
XX
SQ Sequence 380 AA;
XX
QY 1 SYIEVTEGTEATATG 17
    |||||||||||||||
DB 325 SYIEVTEGTEATATG 341

Query Match 100.0%; Score 83; DB 22; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
AA57110
ID AAR57110 standard; peptide; 17 AA.
XX
AC AAR57110;
XX
DT 16-AUG-1994 (first entry)
XX
DE Human megakaryocyte differentiation factor peptide 4.
XX
KW Human megakaryocyte differentiation factor; MDF; thrombopoietin;
KW haematopoietic stimulating factor; thrombocytopoensis; platelet;
KW bone marrow transplantation; cancer chemotherapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "not determined"
FT Misc-difference 13 /note= "not determined"
FT

```

```

FT Misc-difference 16 /note= "not determined"
XX
XX EP583884-A.
PN
XX
PD 23-FEB-1994.
XX
PF 19-JUL-1993; 93EP-0305654.
XX
PR 17-JUL-1992; 92JP-0212305.
PR 04-MAR-1993; 93JP-0067339.
XX
XX (SUNR ) SUNTORY LTD.
PA (TSUJ/) TSUJIMOTO M.
XX
PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
XX
DR WPI: 1994-058782/08.
XX
PT New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
PS Claim 1; Page 19; 47pp; English.
XX
CC Human MDF (see AAR48379) can be isolated from a culture of human
CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
CC stimulates differentiation of megakaryocytes from myeloid cells
CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
CC making it useful for treatment of diseases involving a decrease
CC in platelet number (esp. thrombocytopaenia) such as occurs in bone
CC marrow transplantation and in chemotherapy. MDF has mol.wt.
CC 55-57KD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
CC contains an amino acid sequence comprising at least one of the
CC sequences AAR57107-R57115.
XX
SQ Sequence 17 AA;
XX
QY 2 YIEVTEGTEATATG 17
    |||||||||||||
DB 2 YIEVTEGTEATATG 17

Query Match 83.1%; Score 69; DB 15; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.00025;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 6
AA529329
ID AAY29329 standard; Protein; 217 AA.
XX
AC AAY29329;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human secreted protein clone 1t217_2 protein sequence.
XX
KW Human; secreted protein; nutrition; cytokine; cell proliferation;
KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
KW chemotactic; chemokine; haemostatic; thrombolytic; anti-inflammatory;
KW tumour invasion suppressor; tumour inhibition.
XX
OS Homo sapiens.
XX
PN WO9937674-A1.
XX
PD 29-JUL-1999.
XX
PR 21-JAN-1999; 99WO-US01404.
XX
PR 20-JAN-1999; 99US-0235609.

```

PR 22-JAN-1998; 9805-0072134.
 XX (GEM) GENETICS INST INC.
 XX
 PI Agostino Mf, Clark Hf, Collins-Racie LA, Fechtel K;
 PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Steininger RJ;
 PI Treacy M, Wong Gg;
 XX WPI; 1999-458682/38.
 DR N-PSDB; AAX90444.
 XX
 PT New polynucleotides encoding secreted human proteins derived from,
 PT e.g. fetal brain potentially used as immunostimulators
 PS Claim 22; Page 119-120; 139pp; English.
 XX
 CC The present sequence represents a human secreted protein. Human secreted
 CC protein polynucleotides and proteins are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals, although no
 CC supporting data is given. Suggested activities include nutritional
 CC activity, cytokine and cell proliferation/differentiation activity,
 CC immune stimulating (e.g. as vaccines) or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, and tumour
 CC inhibition activity. The polynucleotides are also stated to be useful
 CC for gene therapy.
 CC
 SQ Sequence 217 AA;
 QY
 Query Match 83.1%; Score 69; DB 20; Length 217;
 Best Local Similarity 76.5%; Pred. No. 0.0049;
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 160 SYLDVSECTEAAATG 176
 1 SYIEVTEGTEATATG 17
 ||:::||||| ||||
 160 SYLDVSECTEAAATG 176
 RESULT 7
 AAU39055
 ID AAU39055 standard; Protein: 217 AA.
 XX
 AC AAU39055;
 XX
 DT 16-JAN-2002 (first entry)
 XX
 DE Human secreted protein It217_2.
 XX
 KW Human; secreted protein; antiinflammatory; immunosuppressive;
 KW neurotropic; neuroprotective; antiarthritic; antimicrobial; vulnery;
 KW cytostatic; antidiabetic; virucide; antinfertility; anticonvulsant;
 KW vasotrophic; antiparkinsonian; immunostimulant; dermatological;
 KW antirheumatic; antitumor; antitumor; osteopathic; tranquiliser;
 KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
 KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
 KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
 KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
 KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
 KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
 KW food supplement; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200175068-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 22-MAR-2001; 2001WO-US09369.
 XX

PR 30-MAR-2000; 2000US-0539330.
 PR 04-DEC-2000; 2000US-0729674.
 XX
 XX (GEM) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
 PI Treacy M, Agostino Mf, Steininger RJ, Spaulding V, Wong Gg;
 PI Clark H, Fechtel K, Merberg D;
 XX WPI; 2001-639363/73.
 DR N-PSDB; AAS99273.
 XX
 PT Secreted human proteins, useful as vaccine for treating various
 PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
 PT nervous system disorders (e.g. stroke)
 PS Disclosure; Page 546-547; 619pp; English.
 XX
 CC The invention relates to novel human secreted proteins, the nucleic
 CC acids encoding them. The protein may exhibit cytokine, cell proliferation
 CC or cell differentiation activity or may induce production of other
 CC cytokines in certain cell populations and may exhibit immune stimulating
 CC or immune suppressing activity, which is useful for the treatment of
 CC various immune deficiencies and disorders e.g. severe combined
 CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
 CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
 CC inflammation. The proteins are also useful in the treatment of diseases
 CC and disorders including tissue, skin and organ transplantation and in
 CC graft-versus-host diseases (GVHD), in the induction of tumour immunity,
 CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
 CC in the treatment of burns, incisions and ulcers; as well as in treatment
 CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
 CC inflammatory processes, diseases of the peripheral nervous system,
 CC Alzheimer's, Parkinson's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
 CC infection of cardiac and central nervous system vessel e.g. stroke,
 CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
 CC protein, having activin- or inhibin-related activities is useful as a
 CC contraceptive based on the ability of inhibins to decrease fertility in
 CC female mammals and decrease spermatogenesis in male mammals. The
 CC proteins and nucleic acids are also useful as food supplements. The
 CC present sequence represents a secreted protein of the invention.
 CC
 SQ Sequence 217 AA;
 QY
 Query Match 83.1%; Score 69; DB 22; Length 217;
 Best Local Similarity 76.5%; Pred. No. 0.0049;
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 160 SYLDVSECTEAAATG 176
 1 SYIEVTEGTEATATG 17
 ||:::||||| ||||
 160 SYLDVSECTEAAATG 176
 RESULT 8
 ABB55764
 ID ABB55764 standard; Protein: 217 AA.
 XX
 AC ABB55764;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 134.
 XX
 KW Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.

XX Homo sapiens.
 OS
 CC
 XX US2001039335-A1.
 PN
 XX
 PD 08-NOV-2001.
 PF
 XX 04-DEC-2000; 2000US-0729674.
 PR
 XX 10-APR-1997; 97US-126425P.
 PR 04-DEC-1997; 97US-067454P.
 PR 20-DEC-1997; 97US-068379P.
 PR 02-JAN-1998; 98US-070346P.
 PR 07-JAN-1998; 98US-070643P.
 PR 08-JAN-1998; 98US-070755P.
 PR 13-JAN-1998; 98US-071304P.
 PR 22-JAN-1998; 98US-072134P.
 PR 30-JAN-1998; 98US-073095P.
 PR 18-FEB-1998; 98US-075038P.
 PR 30-MAR-2000; 2000US-0539330.
 PR 23-NOV-1998; 98US-0197886.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D..
 PA (TREN/) TRENCY M.
 PA (AGOS/) AGOSTINO M J.
 PA (STEL/) STEININGER R J.
 PA (SPAU/) SPAULDING V.
 PA (WONG/) WONG G G.
 PA (CLAR/) CLARK H.
 PA (FECH/) FECHTEL K.
 XX
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steindinger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX
 DR WPI; 2002-040725/05.
 DR N-PSDB; ABA90942.
 XX
 XX New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune
 PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
 PT stroke or inflammations
 XX
 PS Disclosure; Page 274; 349pp; English.
 XX
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 CC ABA90980) and encoded proteins (AB555698-AB555800), especially,
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 CC proteins SEQ ID NO 2 (AB555698) and SEQ ID NO 20 (AB55707) contained in
 CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
 CC are deposited with the American Type Culture Collection (ATCC) with
 CC accession number 98599. The polynucleotides and encoded polypeptides have
 CC cytoskeletal, anti-inflammatory, immunomodulator, vulnerary,
 CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
 CC and anti-inflammatory activity and acting as cytokine modulators,
 CC haematopoietic regulatory, tissue growth modulators and/or cadherin
 CC suppressors. The polypeptides and polynucleotides are useful in gene
 CC therapies, particularly for preventing, treating or ameliorating any of
 CC the following diseases: immune deficiency and disorders; e.g. bacterial
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus
 CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 CC osteoarthritis; central and peripheral nervous system diseases and
 CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
 CC haemophilia, cardiac infarction or stroke; Inflammations, shock, sepsis
 CC or systemic inflammatory response syndrome, ischemia-reperfusion
 CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or

CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
 CC foliaceus.
 CC
 XX
 SQ Sequence 217 AA;
 Query Match 83.1%; Score 69; DB 23; Length 217;
 Best Local Similarity 76.5%; Pred. No. 0.0049;
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SYIEVTEGTEATATG 17
 II:::IIIIII IIII
 Db 160 SYIDVSEEGTEAAATG 176
 RESULT 9
 AAU16928
 ID AAU16928 standard; Protein: 340 AA.
 AC
 XX AAU16928;
 DT 07-NOV-2001 (first entry)
 DE Human novel secreted protein, SEQ ID 169.
 DE
 XX Human: Immunosuppressive; antiarthritic; antirheumatic;
 XX cytoskeletal; cardiast; vasotropic; cerebroprotective; nootropic;
 KW neuroprotective; antiarterial; vitricide; fungicide; ophthalmological;
 KW vulnerary; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.
 OS Homo sapiens.
 XX
 PN WO20015441-A2.
 PN
 XX
 PD 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-0501320.
 PE
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218280.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236127.
 PR 29-SEP-2000; 2000US-0236127.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251890.
 PR 11-DEC-2000; 2000US-0251907.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-476222/51.
 DR N-PSDB; AAS26833.
 XX
 PT Novel polypeptides and polynucleotides useful as diagnostic reagents to
 PT diagnose diseases or disorders associated with aberrant expression or
 PT activity of polypeptides, for treating blood clotting disorder,
 PT haemophilia
 XX
 PS Claim 11; SEQ ID No 169; 601pp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present

Query Match 83.1%; Score 69; DB 22; Length 340;
 Best Local Similarity 76.5%; Pred. No. 0.0082;

[illegible]

PR	31-MAR-1998;	98US-0080167.
PR	15-MAY-1998;	98US-0085720.
XX		
PA	(UROG-) UROGENESYS INC.	
PA	(AFAR/) AFAR D E.	
PA	(HUBE/) HUBERT R S.	
PA	(LEON/) LEONG K.	
PA	(RAIT/) RAITANO A B.	
PA	(SAFE/) SAFFRAN D C.	
XX		
PI	Afar DE, Hubert RS, Leong K, Raitano AB, Safran DC;	
DR	WPI: 2000-039068/03.	
DR	N-PSDB; AA223049.	
XX		
PT	New isolated prostaticin gene, used to develop products for, e.g.	
PT	treating and determining susceptibility to prostate cancer	
XX		
PS	Claim 1; Fig 1; 63pp; English.	
XX		
CC	The present sequence is that of a PROSTAPIN protein which is a member of	
CC	serine protease inhibitor (serpin) family and expressed in normal	
CC	prostate and locally confined prostate cancer. The sequence is derived	
CC	from overlapping sequences of human cDNA clones 11P2a6 and 103. PROSTAPIN	
CC	contains a reactive-site loop characteristic of serpin family 1n	
CC	the carboxy-terminal region. The protein may function as a prostate-	
CC	specific tumour suppressor, apoptosis-inducer or apoptosis-modulator.	
CC	The sequence is useful in diagnosing, prognosing, monitoring and	
CC	treating prostate cancer, particularly advanced stage and metastatic	
CC	prostate cancer.	
XX		
SO	Sequence 379 AA;	
	Query Match 83.1%; Score 69; DB 21; Length 379;	
	Best Local Similarity 76.5%; Pred. No. 0.0093;	
	Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	1 SYIEVTEGTEPATATG 17	
	:	
Db	322 SYLDVSEGTAAATG 338	
RESULT 14		
AAV15156		
ID	AAV15156 standard; Protein; 617 AA.	
XX		
AC	AAV15156;	
XX		
DT	07-FEB-2000 (first entry)	
XX		
DE	Human prostate serine protease inhibitor (PROSTAPIN) mutant.	
KN	PROSTAPIN mutant; serine protease inhibitor; serpin family; prostate;	
KN	LARC-9 AD xenograft; androgen dependent LARC-9 xenograft; cDNA clone 2;	
KN	diagnosis; prognosis; treatment; prostate cancer; human;	
KN	metastatic prostate cancer.	
XX		
OS	Homo sapiens.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 90 /note= "Corresponds to stop codon resulting due to	
FT	point mutation"	
FT	Misc-difference 215 /note= "Corresponds to stop codon"	
FT	Misc-difference 225 /note= "Corresponds to stop codon"	
FT	Misc-difference 232 /note= "Corresponds to stop codon"	
FT	Misc-difference 233 /note= "Corresponds to stop codon"	
FT	Misc-difference 289 /note= "Corresponds to stop codon"	
FT		

FT MISC-difference 309 /note- "Corresponds to stop codon"
 FT MISC-difference 318 /note- "Corresponds to stop codon"
 FT MISC-difference 322 /note- "Corresponds to stop codon"
 FT MISC-difference 331 /note- "Corresponds to stop codon"
 FT MISC-difference 332 /note- "Corresponds to stop codon"
 FT MISC-difference 367 /note- "Corresponds to stop codon"
 FT MISC-difference 369 /note- "Corresponds to stop codon"
 FT MISC-difference 374 /note- "Corresponds to stop codon"
 FT MISC-difference 378 /note- "Corresponds to stop codon"
 FT MISC-difference 425 /note- "Corresponds to stop codon"
 FT MISC-difference 430 /note- "Corresponds to stop codon"
 FT MISC-difference 148 /note- "Corresponds to stop codon"
 FT MISC-difference 181 /note- "Wild type Met substituted by Thr"
 FT MISC-difference 188 /note- "Wild type Thr substituted by Ala"
 FT MISC-difference /note- "Wild type Arg substituted by Trp"
 FT Region 194.432 /note- "Insertion sequence corresponding to unspliced intron in the coding sequence"
 FT MISC-difference 518 /note- "Wild type Ile substituted by Thr"
 FT MISC-difference 528 /note- "Wild type Pro substituted by Ser"
 XX WO958560-A2.
 PD 18-NOV-1999.
 XX 31-MAR-1999; 99WO-US07123.
 XX 31-MAR-1998; 98US-0080167.
 PR 15-MAY-1998; 98US-0085720.
 XX (UROG-) UROGENESYS INC.
 PA (AFAR/) AFAR D E.
 PA (HUBE/) HUBERT R S.
 PA (LEON/) LEONG K.
 PA (RAIT/) RAITANO A B.
 PA (SAFF/) SAFFRAN D C.
 XX Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
 PI WPI: 2000-039068/03.
 DR N-PSDB: AA229050.
 XX New isolated prostate gene, used to develop products for, e.g.
 PT treating and determining susceptibility to prostate cancer.
 XX Example 5; Fig 6; 63pp: English.
 XX The present sequence is a PROSTAPIN (prostate serine protease inhibitor)
 CC mutant encoded by cDNA clone 2 derived from cDNA library of
 CC prostate cancer xenograft LAPC-9 AD which was generated from a bone
 CC tumour biopsy of a patient with hormone-refractory metastatic prostate
 CC cancer. Clone 2 is a partially unspliced version of wild type PROSTAPIN
 CC coding sequence with an insertion sequence and several point mutations.
 CC PROSTAPIN is expressed in normal prostate and locally confined prostate
 CC cancer, while its expression is lost or attenuated in advanced stage
 CC cancer. Wild type PROSTAPIN functionality is either altered or completely
 CC lost in the mutant. PROSTAPIN is useful in diagnosing, prognosing,
 CC monitoring and treating prostate cancer, particularly advanced stage and

CC metastatic prostate cancer.
 XX Sequence 617 AA;
 SO Query Match 83.1%; Score 69; DB 21; Length 617;
 Best Local Similarity 76.5%; Pred. NO. 0.016;
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SYIEVTEGTEATATG 17
 DB 560 SYLDVSEGTAAATG 576
 RESULT 15
 AAAY72661
 ID AA72661 standard; Protein: 135 AA.
 XX AAAY72661;
 AC 31-MAY-2001 (first entry)
 DT
 XX Human headpin (for head and neck serpin) partial protein #2.
 DE Human; headpin; head and neck serpin; serine protease inhibitor;
 KW cytosolic; chromosome 18q21.3-22; therapy; cancer.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH MISC-difference 1..2 /note- "Encoded by ATA"
 FT
 XX WO200116324-A2.
 XX 08-MAR-2001.
 XX 31-AUG-2000; 2000WO-US24214.
 XX 31-AUG-1999; 99US-0151776.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA Clayman GL, Nakashima T, Spring PM;
 DR WPI: 2001-169011/17.
 DR N-PSDB: AAD02741.
 XX New isolated polynucleotide encoding a headpin (head and neck serpin
 PT (serine protease inhibitor)) for diagnosing and treating cancers and
 PT for identifying substances that modulate the headpin activity -
 XX Claim 21; Page 210; 213pp: English.
 XX The present sequence is human headpin (for head and neck serpin (serine
 CC protease inhibitor)) partial protein. Headpin is a differentially
 CC expressed serine protease inhibitor that belongs to the ov-serpin
 CC family. It has been grouped into the cluster of serpins located at
 CC chromosome 18q21.3/18q22. The headpin cDNA and protein are used to
 CC diagnose and treat cancer, such as of the head and neck, brain, lung,
 CC liver, spleen, kidney, lymph node, small intestine, pancreas, blood
 CC cells, colon, stomach, breast, endometrium, prostate, testicle, ovary,
 CC skin, oesophagus, bone marrow or blood. They are used to identify
 CC substances that modulate the headpin activity. Antibodies of headpin
 CC are also used to diagnose cancers using e.g. enzyme linked immunosorbent
 CC assays (ELISA).
 XX Sequence 135 AA;
 SO Query Match 80.7%; Score 67; DB 22; Length 135;
 Best Local Similarity 76.5%; Pred. NO. 0.0058;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SYIEVTEGTEATATG 17

Fri Jul 11 15:10:57 2003

us-10-091-442-4.rag

Page 10

Db 78 SEVAVTEEGTEAAATG 94

Search completed: July 11, 2003, 11:54:20
Job time : 10.7547 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 3.0395 Seconds
(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-4

Perfect score: 83

Sequence: 1 SYIEVTEGTEATATG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	17	US-08-611-977-4	Sequence 4, Appl1
2	83	100.0	380	US-08-472-659-34	Sequence 34, Appl1
3	83	100.0	380	US-08-474-661-34	Sequence 34, Appl1
4	83	100.0	380	US-08-611-977-34	Sequence 34, Appl1
5	69	83.1	17	US-08-472-659-4	Sequence 4, Appl1
6	69	83.1	17	US-08-474-661-4	Sequence 4, Appl1
7	67	80.7	391	US-09-123-912-110	Sequence 110, App
8	67	80.7	391	US-09-643-597-110	Sequence 110, App
9	67	80.7	400	US-09-123-912-112	Sequence 112, App
10	67	80.7	400	US-09-643-597-112	Sequence 112, App
11	62	74.7	376	US-09-200-965-2	Sequence 2, Appl1
12	62	74.7	379	US-08-121-714-4	Sequence 4, Appl1
13	62	74.7	379	US-08-477-108A-4	Sequence 4, Appl1
14	62	74.7	379	US-08-477-112-4	Sequence 4, Appl1
15	62	74.7	379	PCT-US93-08322-4	Sequence 4, Appl1
16	61	73.5	376	US-08-745-995A-12	Sequence 12, Appl1
17	61	73.5	376	US-08-745-995A-24	Sequence 24, Appl1
18	61	73.5	376	US-08-745-995A-30	Sequence 30, Appl1
19	61	73.5	393	US-08-745-995A-20	Sequence 20, Appl1
20	61	73.5	398	US-08-745-995A-26	Sequence 26, Appl1
21	61	73.5	399	US-08-745-995A-8	Sequence 8, Appl1
22	59	71.1	390	US-08-568-147B-2	Sequence 2, Appl1
23	59	71.1	390	US-09-266-910-3	Sequence 3, Appl1
24	59	71.1	390	US-09-266-910-4	Sequence 4, Appl1
25	58	69.9	374	US-08-464-148-2	Sequence 2, Appl1
26	58	69.9	374	US-08-385-500-2	Sequence 2, Appl1
27	58	69.9	374	US-08-846-784-2	Sequence 2, Appl1

28	58	69.9	376	1	US-08-464-148-4	Sequence 4, Appl1
29	58	69.9	376	1	US-08-385-500-4	Sequence 4, Appl1
30	58	69.9	376	1	US-08-846-784-4	Sequence 4, Appl1
31	56	67.5	375	1	US-08-121-714-8	Sequence 8, Appl1
32	56	67.5	375	1	US-08-477-108A-8	Sequence 8, Appl1
33	56	67.5	375	2	US-08-477-112-8	Sequence 8, Appl1
34	56	67.5	375	5	PCT-US93-08322-8	Sequence 8, Appl1
35	56	67.5	407	3	US-08-948-997-3	Sequence 3, Appl1
36	56	67.5	410	4	US-09-348-817A-3	Sequence 3, Appl1
37	55	66.3	376	4	US-08-745-995A-6	Sequence 6, Appl1
38	55	66.3	376	4	US-08-745-995A-36	Sequence 36, Appl1
39	55	66.3	390	4	US-08-745-995A-18	Sequence 18, Appl1
40	55	66.3	397	4	US-08-745-995A-2	Sequence 2, Appl1
41	55	66.3	397	4	US-08-745-995A-32	Sequence 32, Appl1
42	55	66.3	410	3	US-08-948-997-2	Sequence 3, Appl1
43	55	66.3	410	4	US-09-348-817A-2	Sequence 2, Appl1
44	55	66.3	420	4	US-08-745-995A-14	Sequence 14, Appl1
45	54	65.1	382	1	US-07-768-286B-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-611-977-4
: Sequence 4, Application US/08611977
: Patent No. 5972886
GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masafumi
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSUROOKA, No. 5972886uo
: APPLICANT: NAKAZATO, Hiroshi
: APPLICANT: MIURA, Kenju
: APPLICANT: ISHIDA, No. 5972886uhitro
: APPLICANT: KURIHARA, Tatsuya
: APPLICANT: YAMAICHI, Kozo
: APPLICANT: YAMAGUCHI, No. 5972886om1
: TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Burns, Doane, Swecker & Mathis
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/611.977
: FILING DATE: 06-MAR-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/091,028
: FILING DATE: 14-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 4-212305
: FILING DATE: 17-JUL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-067339
: FILING DATE: 04-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: McGowan, Malcolm K.
: REGISTRATION NUMBER: 39,300
: REFERENCE/DOCKET NUMBER: 001560-204
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-4

Query Match 100.0%; Score 83; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYIEVTEEGTEAATG 17
DB 1 SYIEVTEEGTEAATG 17

RESULT 2

US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-472-659-34

Query Match

100.0%; Score 83; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYIEVTEEGTEAATG 17
DB 325 SYIEVTEEGTEAATG 341

RESULT 3

US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 83; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYIEVTEEGTEAATG 17
DB 325 SYIEVTEEGTEAATG 341

RESULT 4

US-08-611-977-34
: Sequence 34, Application US/08611977
: Patent No. 5972886
: GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masafumi
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSURUOKA, No. 5972886uo
: APPLICANT: NAKAZATO, Hiroshi
: APPLICANT: MIURA, Kenju
: APPLICANT: ISHIDA, No. 5972886uhiro
: APPLICANT: KURIHARA, Tatsuya
: APPLICANT: YAMAGUCHI, Kozo
: APPLICANT: YAMAGUCHI, No. 5972886om1
: TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Burns, Doane, Swecker & Mathis
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/611,977
: FILING DATE: 06-MAR-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/091,028
: FILING DATE: 14-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 4-212305
: FILING DATE: 17-JUL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-067339
: FILING DATE: 04-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: McGowan, Malcolm K.
: REGISTRATION NUMBER: 39,300
: REFERENCE/DOCKET NUMBER: 001560-204
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 34:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 380 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-611-977-34

Query Match 100.0%; Score 83; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 6,4e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
DB 325 SYIEVTEGTEATATG 341

RESULT 5
US-08-472-659-4
: Sequence 4, Application US/08472659
: Patent No. 5831030
: GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masafumi
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSURUOKA, No. 5831030uo
: APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-472-659-4

Query Match 83.1%; Score 69; DB 2; Length 17;
Best Local Similarity 87.5%; Pred. No. 3,2e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YIEVTEGTEATATG 17
DB 2 YIEVTEGTEATATG 17

RESULT 6
US-08-474-661-4
: Sequence 4, Application US/08474661
: Patent No. 5874253
: GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masafumi
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSURUOKA, No. 5874253uo
: APPLICANT: NAKAZATO, Hiroshi
: APPLICANT: MIURA, Kenju
: APPLICANT: ISHIDA, No. 5874253uhiro
: APPLICANT: KURIHARA, Tatsuya
: APPLICANT: YAMAGUCHI, Kozo
: APPLICANT: YAMAGUCHI, No. 5874253om1
: TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
: NUMBER OF SEQUENCES: 34

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-661-4

Query Match      83.1%; Score 69; DB 2; Length 17;
Best Local Similarity 87.5%; Pred. No. 3.2e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 YIEVTEEGTEATATG 17
DB      2 YIEVTEEGTEAAXAG 17

RESULT 7
US-09-123-912-110
; Sequence 110, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 110
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-123-912-110

Query Match      80.7%; Score 67; DB 4; Length 391;
Best Local Similarity 76.5%; Pred. No. 0.0027;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 SYIEVTEEGTEATATG 17
DB      334 SEVAVTEEGTEAAXAG 350

RESULT 8
US-09-643-597-110
; Sequence 110, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-643-597-110

Query Match      80.7%; Score 67; DB 4; Length 391;
Best Local Similarity 76.5%; Pred. No. 0.0027;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 SYIEVTEEGTEATATG 17
DB      334 SEVAVTEEGTEAAXAG 350

RESULT 9
US-09-123-912-112
; Sequence 112, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 112
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-123-912-112

Query Match      80.7%; Score 67; DB 4; Length 400;
Best Local Similarity 76.5%; Pred. No. 0.0028;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```

: TELEFAX: 201-994-1744
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 376 AMINO ACIDS
: TYPE: AMINO ACID.
: STRANDEDNESS:
: TOPOLOGY: LINEAR
: MOLECULE TYPE: PROTEIN
: US-09-200-965-2

Query Match
Best Local Similarity 74.7%; Score 62; DB 4; Length 376;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAAT 16
: : : : : : : : : : : : : : : :
Db 319 SFVEVNEGTAAAT 334

RESULT 12
US-08-121-714-4
: Sequence 4, Application US/08121714
: Patent No. 5470970
: GENERAL INFORMATION:
: APPLICANT: Sager, Ruth
: TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
: TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: COMPUTER: IBM PS/2 Model 502 or 535x
: OPERATING SYSTEM: MS-DOS (Version 5.0)
: SOFTWARE: WordPerfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/121,714
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/938,823
: FILING DATE: 09/01/92
: APPLICATION NUMBER: 07/844,296
: FILING DATE: 02/28/92
: APPLICATION NUMBER: 07/662,216
: FILING DATE: 02/28/91
: ATTORNEY/AGENT INFORMATION:
: NAME: Fraser, Janis K.
: REGISTRATION NUMBER: 34,819
: REFERENCE/DOCKET NUMBER: 00530/072001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 379
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: US-08-121-714-4

Query Match
Best Local Similarity 74.7%; Score 62; DB 1; Length 379;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAAT 16
: : : : : : : : : : : : : : : :

```

Db 322 SEVEVEEGTEAAAT 337

```
RESULT 13
US-08-477-108A-4
; Sequence 4, Application US/08477108A
; Patent No. 5801001
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Zou, Zhigang
; APPLICANT: Anisowicz, Anthony
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,108A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ. ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-477-108A-4

Query Match 74.7%; Score 62; DB 1; Length 379;
Best Local Similarity 75.0%; Pred. No. 0.017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 SYIEVEEGTEAAAT 16
Db 322 SEVEVEEGTEAAAT 337
```

```
RESULT 14
US-08-477-112-4
; Sequence 4, Application US/08477112
; Patent No. 5905023
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,112
; FILING DATE:
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ. ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-477-112-4

Query Match 74.7%; Score 62; DB 2; Length 379;
Best Local Similarity 75.0%; Pred. No. 0.017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 SYIEVEEGTEAAAT 16
Db 322 SEVEVEEGTEAAAT 337
```

```
RESULT 15
PCT-US93-08322-4
; Sequence 4, Application PC/TUS9308322
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 555X
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08322
; FILING DATE:
; CLASSIFICATION:
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 379
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-08322-4

Query Match 74.7%; Score 62; DB 5; Length 379;
Best Local Similarity 75.0%; Pred. No. 0.017;
Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYIEVTEEGTEATAT 16
I::I I I I I I I I I
Db 322 SPVEVNEEGTEAAAT 337

Search completed: July 11, 2003, 12:02:07
Job time : 4.0395 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 4.98337 Seconds
(Without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-4

Perfect score: 83
Sequence: 1 SYIEVTEEGTEATATG 17

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_NEM_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCr_NEM_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/PCrUS_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US08_NEM_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	83	100.0	17	US-10-091-442-4	Sequence 4, Appl1
2	83	100.0	17	US-09-140-719-4	Sequence 4, Appl1
3	83	100.0	380	US-10-091-442-34	Sequence 34, Appl1
4	83	100.0	380	US-09-140-719-34	Sequence 34, Appl1
5	69	83.1	217	US-09-729-674-134	Sequence 134, Appl1
6	69	83.1	340	US-09-764-898-169	Sequence 169, Appl1
7	69	83.1	379	US-10-012-896-946	Sequence 946, Appl1
8	69	83.1	379	US-09-895-814-946	Sequence 946, Appl1
9	69	83.1	392	US-10-198-070-62	Sequence 62, Appl1
10	69	83.1	392	US-10-198-070-67	Sequence 67, Appl1
11	69	83.1	392	US-10-198-070-70	Sequence 70, Appl1
12	69	83.1	617	US-10-012-896-947	Sequence 947, Appl1
13	69	83.1	617	US-09-895-814-947	Sequence 947, Appl1
14	67	80.7	391	US-10-157-031-355	Sequence 355, Appl1
15	67	80.7	391	US-09-466-396A-110	Sequence 110, Appl1
16	67	80.7	391	US-09-735-705-110	Sequence 110, Appl1
17	67	80.7	391	US-09-850-716A-110	Sequence 110, Appl1
18	67	80.7	391	US-09-897-778-110	Sequence 110, Appl1
19	67	80.7	400	US-09-466-396A-112	Sequence 112, Appl1

20	67	80.7	400	US-09-735-705-112	Sequence 112, Appl1
21	67	80.7	400	US-09-850-716A-112	Sequence 112, Appl1
22	67	80.7	400	US-09-897-778-112	Sequence 112, Appl1
23	62	74.7	437	US-09-925-301-1200	Sequence 1200, Appl1
24	62	74.7	439	US-09-925-300-1440	Sequence 1440, Appl1
25	59	71.1	377	US-10-165-605A-27	Sequence 27, Appl1
26	59	71.1	377	US-09-823-187-31	Sequence 27, Appl1
27	59	71.1	377	US-09-910-430-27	Sequence 27, Appl1
28	59	71.1	390	US-08-731-566-2	Sequence 2, Appl1
29	59	71.1	390	US-08-823-187-27	Sequence 27, Appl1
30	59	71.1	390	US-08-823-187-28	Sequence 28, Appl1
31	59	71.1	390	US-09-823-187-29	Sequence 29, Appl1
32	59	71.1	390	US-09-823-187-30	Sequence 30, Appl1
33	58	69.9	374	US-09-974-298-95	Sequence 95, Appl1
34	57	68.7	417	US-10-174-550-452	Sequence 452, Appl1
35	57	68.7	417	US-10-176-758-452	Sequence 452, Appl1
36	57	68.7	417	US-10-175-737-452	Sequence 452, Appl1
37	57	68.7	417	US-10-173-706-452	Sequence 452, Appl1
38	57	68.7	417	US-10-175-738-452	Sequence 452, Appl1
39	57	68.7	417	US-10-175-752-452	Sequence 452, Appl1
40	57	68.7	417	US-10-176-482-452	Sequence 452, Appl1
41	57	68.7	417	US-10-176-757-452	Sequence 452, Appl1
42	57	68.7	417	US-10-176-757-452	Sequence 452, Appl1
43	57	68.7	417	US-10-176-913-452	Sequence 452, Appl1
44	57	68.7	417	US-10-180-552-452	Sequence 452, Appl1
45	57	68.7	417	US-10-180-557-452	Sequence 452, Appl1
				US-10-173-700-452	Sequence 452, Appl1

ALIGNMENTS

RESULT 1
US-10-091-442-4
Sequence 4, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSURUOKA, NO. US20020164711A1no
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, No. US20020164711A1uhiro
KURIHARA, Tatsuya
YAMACHI, Kozo
YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091.442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140.719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474.661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091.028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-091-442-4

Query Match 100.0%; Score 83; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYEVTEEGTEATATG 17
Db 1 SYEVTEEGTEATATG 17

RESULT 2
US-09-140-719-4
Sequence 4, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUROOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-4

Query Match 100.0%; Score 83; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYEVTEEGTEATATG 17
Db 1 SYEVTEEGTEATATG 17

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUROOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-MAR-2002
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 380 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 83; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEAATG 17
DB 325 SYIEVTEEGTEAATG 341

RESULT 4
US-09-140-719-34
; Sequence 34, Application US/09140719
; Patent No. US20010026931A1
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IMASA, Fuyuki
; APPLICANT: TSUBOUOKA, No. US20010026931A1uo
; APPLICANT: MAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. US20010026931A1uhro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. US20010026931A1oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid

;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 83; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEAATG 17
DB 325 SYIEVTEEGTEAATG 341

RESULT 5
US-09-729-674-134
; Sequence 134, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectchel, Kim
; APPLICANT: Genetics-Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6035-64X
; CURRENT APPLICATION NUMBER: US/09/729,674
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-674-134

Query Match 83.1%; Score 69; DB 10; Length 217;
Best Local Similarity 76.5%; Pred. No. 0.0014;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEAATG 17
DB 160 SYLDVSEEGTEAATG 176

RESULT 6
US-09-764-898-169
; Sequence 169, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, proteins, and antibodies
; FILE REFERENCE: P0201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 169
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-169

Query Match 83.1%; Score 69; DB 10; Length 340;
Best Local Similarity 76.5%; Pred. No. 0.0023;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEECTEATATG 17
||:::||||| ||||
Db 283 SYLDVSECTEAAATG 299

RESULT 7

US-10-012-896-946

; Sequence 946, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-946

Query Match 83.1%; Score 69; DB 9; Length 379;
Best Local Similarity 76.5%; Pred. No. 0.0026;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEECTEATATG 17
||:::||||| ||||
Db 322 SYLDVSECTEAAATG 338

RESULT 8

US-09-895-814-946

; Sequence 946, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-946

Query Match 83.1%; Score 69; DB 9; Length 379;
Best Local Similarity 76.5%; Pred. No. 0.0026;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEECTEATATG 17
||:::||||| ||||
Db 322 SYLDVSECTEAAATG 338

RESULT 9

US-10-198-070-62

; Sequence 62, Application US/10198070
; Publication No. US20030109437A1
; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-62

Query Match 83.1%; Score 69; DB 9; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0027;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEECTEATATG 17
||:::||||| ||||
Db 335 SYLDVSECTEAAATG 351

RESULT 10
US-10-198-070-67
: Sequence 67, Application US/10198070
: Publication No. US20030109437A1
: GENERAL INFORMATION:
: APPLICANT: AVERBACK, PAUL
: APPLICANT: GEMMELL, JACK
: TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
: TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 59003.000008
: CURRENT APPLICATION NUMBER: US/10/198.070
: CURRENT FILING DATE: 2002-07-19
: PRIOR APPLICATION NUMBER: 60/306,161
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/306,150
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/331,477
: PRIOR FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 67
: LENGTH: 392
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-198-070-67

Query Match 83.1%, Score 69; DB 9; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0027;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYIVTEGTEATATG 17
||:|||||
Db 335 SYLDVSEGTAAATG 351

RESULT 11
US-10-198-070-70
: Sequence 70, Application US/10198070
: Publication No. US20030109437A1
: GENERAL INFORMATION:
: APPLICANT: AVERBACK, PAUL
: APPLICANT: GEMMELL, JACK
: TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
: TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 59003.000008
: CURRENT APPLICATION NUMBER: US/10/198.070
: CURRENT FILING DATE: 2002-07-19
: PRIOR APPLICATION NUMBER: 60/306,161
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/306,150
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/331,477
: PRIOR FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 70
: LENGTH: 392
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-198-070-70

Query Match 83.1%, Score 69; DB 9; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0027;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYIVTEGTEATATG 17
||:|||||
Db 335 SYLDVSEGTAAATG 351

RESULT 12

US-10-012-896-947
: Sequence 947, Application US/10012896
: Publication No. US20020183251A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yuguu
: APPLICANT: Kalos, Michael D.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stolck, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aljun
: APPLICANT: Skeiky, Yasir A.W.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Hural, John
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Vinals de Baassols, Carlota
: APPLICANT: Foy, Teresa
: APPLICANT: Fanger, Gary R.
: APPLICANT: Mantanabe, Yoshihiro
: APPLICANT: Meagher, Madeline Joy
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
: FILE REFERENCE: 210121.427C27
: CURRENT APPLICATION NUMBER: US/10/012.896
: CURRENT FILING DATE: 2001-12-10
: NUMBER OF SEQ ID NOS: 1011
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 947
: LENGTH: 617
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: 90, 215, 225, 232, 233, 289, 309, 318, 322, 331, 332, 367,
: LOCATION: 369, 374, 378, 425, 430
: OTHER INFORMATION: Xaa - Any Amino Acid
US-10-012-896-947

Query Match 83.1%, Score 69; DB 9; Length 617;
Best Local Similarity 76.5%; Pred. No. 0.0045;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYIVTEGTEATATG 17
||:|||||
Db 560 SYLDVSEGTAAATG 576

RESULT 13
US-09-895-814-947
: Sequence 947, Application US/09895814
: Publication No. US20020193296A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yuguu
: APPLICANT: Kalos, Michael D.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stolck, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aljun

```

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepier, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 947
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)..(617)
; OTHER INFORMATION: Xaa - Any amino acid
US-09-895-814-947

```

```

Query Match      83.1%; Score 69; DB 9; Length 617;
Best Local Similarity 76.5%; Pred. No. 0.0045;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 SYIEVTEGTEATATG 17
        11::1111111111
DB      560 SYLDVSECTEAAATG 576

```

```

RESULT 14
US-10-157-031-355
; Sequence 355, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krivoskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 355
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-355

```

```

Query Match      80.7%; Score 67; DB 9; Length 391;
Best Local Similarity 76.5%; Pred. No. 0.0057;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 SYIEVTEGTEATATG 17
        1::1111111111
DB      334 SFVAVTEGTEAAATG 350

```

```

RESULT 15
US-09-466-396A-110
; Sequence 110, Application US/09466396A
; Publication No. US20030119763A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND

```

```

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-466-396A-110

```

```

Query Match      80.7%; Score 67; DB 9; Length 391;
Best Local Similarity 76.5%; Pred. No. 0.0057;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 SYIEVTEGTEATATG 17
        1::1111111111
DB      334 SFVAVTEGTEAAATG 350

```

```

Search completed: July 11, 2003, 12:37:33
Job time : 5.98337 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 3.42827 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-4

Perfect score: 83
Sequence: 1 SYIEVTEGTEAATG 17

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	80.7	391	2 JC7118	headpin serine pro
2	67	74.7	376	1 A48681	placental thrombin
3	62	74.7	376	2 S38962	serpin - pig
4	60	72.3	379	2 S27383	elastase inhibitor
5	60	72.3	388	1 DYCH	ovalbumin-related
6	59	72.3	399	1 DKBH2	protein 24 - barley
7	59	71.1	378	2 A57488	proteinase inhibitor
8	59	71.1	390	2 I38201	squamous cell carc
9	58	71.1	390	2 I38202	leupin precursor -
10	58	69.9	374	2 A59273	proteinase inhibitor
11	58	69.9	376	2 B59273	proteinase inhibitor
12	58	69.9	397	2 T06183	serpin - barley
13	57	68.7	374	2 A41903	hypothetical prote
14	57	68.7	398	2 S39819	serpin - barley
15	56	67.5	213	2 A26423	serine proteinase
16	56	67.5	379	2 A42421	leukocyte elastase
17	56	67.5	410	2 S70647	neuroserpin precu
18	55	67.5	418	2 S31507	serine proteinase
19	55	66.3	403	1 S08102	serine proteinase
20	54	65.1	407	2 D84518	probable serpin (i
21	54	65.1	415	2 A32853	plasmaogen activa
22	54	65.1	559	2 B96648	hypothetical prote
23	53	63.9	397	2 I39184	homapin - human
24	53	63.9	398	2 S65782	serpin - wheat
25	53	63.9	399	2 T06488	serpin WZS2 - whea
26	53	63.9	429	2 S49585	probable serine pr
27	52	62.7	93	2 S66290	alpha 1 antichymot
28	52	62.7	222	1 WNVZ84	antithrombin-III h
29	52	62.7	252	2 S66289	alpha 1 antichymot

30	52	62.7	341	1 WNVZHI	antithrombin-III h
31	52	62.7	344	2 S46871	B1R protein - vari
32	52	62.7	344	2 E72173	D2R protein - vari
33	52	62.7	344	2 T28611	hypothetical prote
34	52	62.7	345	1 WNVZM2	antithrombin-III h
35	52	62.7	391	2 H96517	protein T2E6.22 (1
36	51	61.4	64	2 S15635	serpin Spi2 (clone
37	51	61.4	64	2 S15633	serpin Spi2 (clone
38	51	61.4	66	2 S15629	serpin Spi2 (clone
39	51	61.4	216	2 B26422	serine proteinase
40	51	61.4	344	1 S24676	serpin - Ectromeli
41	51	61.4	353	1 WNVZS1	antithrombin-III h
42	51	61.4	353	1 WNVZC2	antithrombin-III h
43	51	61.4	366	2 T16119	hypothetical prote
44	51	61.4	369	2 JH0493	alpha-1-antichymot
45	51	61.4	397	2 I48717	proteinase inhibitor

ALIGNMENTS

```

RESULT 1
JC7118
headpin serine proteinase inhibitor - human
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 26-May-2000
C:Accession: JC7118
R:Spring, P.; Nakashima, T.; Frederick, M.; Henderson, Y.; Clayman, G.
Biochem. Biophys. Res. Commun. 264, 299-304, 1999
A:Title: Identification and CDNA cloning of headpin, a novel differentially expressed
A:Reference number: JC7118; MUID:99458661; PMID:10527881
A:Accession: JC7118
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-391 <SPR>
A:Cross-references: GB:A5169949; NID:95911368; PIDN:A055765.1; PID:95911369
C:Genetics:
A:Map position: 18q21.3-18q22
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor
F:355/Inhibitory site: Val (unidentified proteinase) #status predicted

Query Match
Best Local Similarity 80.7%; Score 67; DB 2; Length 391;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY      1 SYIEVTEGTEAATG 17
      11: ||||| |||
Db      334 SFVAVTEGTEAATG 350

RESULT 2
A48681
placental thrombin inhibitor - human
N:Alternate names: cytoplasmic antiproteinase; intracellular serine proteinase inhibi
C:Species: Homo sapiens (man)
C:Date: 07-Apr-1994 #sequence_revision 07-Jul-1995 #text_change 18-Jun-1999
C:Accession: A48681; A54352; A46672; C46672; S35750
R:Coughlin, P.; Sun, J.; Cerruti, L.; Salem, H.H.; Bird, P.
Proc. Natl. Acad. Sci. U.S.A. 90, 9417-9421, 1993
A:Title: Cloning and molecular characterization of a human intracellular serine prote
A:Reference number: A48681; MUID:94022286; PMID:8415716
A:Accession: A48681
A:Molecule type: mRNA
A:Residues: 1-376 <CDU>
A:Cross-references: GB:222658; NID:9297411; PIDN:CAA80373.1; PID:9297412
A:Experimental source: placenta
A>Note: authors translated the codon GAA for residue 198 as Gly
R:Morgenstern, K.A.; Sprecher, C.; Holth, L.; Foster, D.; Grant, F.J.; Ching, A.; Kis
Biochemistry 33, 3432-3441, 1994
A:Title: Complementary DNA cloning and kinetic characterization of a novel intracellu
A:Reference number: A54352; MUID:94183847; PMID:8136380
A:Accession: A54352

```

A:Molecule type: mRNA
 A:Residues: 1-174, 'E', 176-361, 'S', 363-376 <MOR>
 A:Cross-references: GB:S69272; NID:9546087; PIDN:AA30320.1; PID:9546088
 A:Experimental source: placenta
 A>Note: sequence extracted from NCBI backbone (NCBI:145231, NCBI:145232)
 R:Coughlin, P.B.; Tetaz, T.; Salem, H.H.
 J. Biol. Chem. 268, 9341-9347, 1993
 A:Title: Identification and purification of a novel serine proteinase inhibitor.
 A:Reference number: A46672; MUID:93252826; PMID:8486644
 A:Accession: A46672
 A:Molecule type: protein
 A:Residues: 47-60;63-81;91-98 <CO2>
 A:Experimental source: placenta, leukemic cell line K562
 A>Note: sequence modified after extraction from NCBI backbone
 C:Genetics:
 A:Gene: GDB:P16
 A:Cross-references: GDB:252025; OMIM:173321
 A:Map position: 6p25-6p24.3
 C:Superfamily: antithrombin III
 C:Keywords: blocked amino end; cytosol; serine proteinase inhibitor
 F:341/inhibitory site: Arg (thrombin) #status predicted

Query Match 74.7%; Score 62; DB 1; Length 376;
 Best Local Similarity 75.0%; Pred. No. 0.015;
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAT 16
 I::||| ||||| |||
 Db 319 SFVEVNEGTEAAAT 334

RESULT 3
 538962
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Jul-1998
 C:Accession: S38962
 R:Reschauer, W.F.; Mentele, R.; Sommerhoff, C.P.
 Eur. J. Biochem. 217, 519-526, 1993
 A:Title: Primary structure of a porcine leukocyte serpin.
 A:Reference number: S38962; MUID:94039085; PMID:7901009
 A:Accession: S38962
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-378 <RES>
 A>Note: the sequence from Fig. 6 is inconsistent with that from Fig. 5 in having 256-Asn
 C:Superfamily: antithrombin III

Query Match 74.7%; Score 62; DB 2; Length 378;
 Best Local Similarity 75.0%; Pred. No. 0.015;
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAT 16
 I::||| ||||| |||
 Db 321 SFVEVNEGTEAAAT 336

RESULT 4
 527383
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Sep-1997
 C:Accession: S27383; S65750
 R:Remold-O'Donnell, E.; Chin, J.; Alberts, M.
 Proc. Natl. Acad. Sci. U.S.A. 89, 5635-5639, 1992
 A:Title: Sequence and molecular characterization of human monocyte/neutrophil elastase
 A:Reference number: S27383; MUID:92302296; PMID:1376927
 A:Accession: S27383
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-379 <REN>
 A:Cross-references: EMBL:M93056
 R:Packard, B.Z.; Lee, S.S.; Remold-O'Donnell, E.; Komoriya, A.

Biochim. Biophys. Acta 1269, 41-50, 1995
 A:Title: A serpin from human tumor cells with direct lymphoid immunomodulatory activity
 A:Reference number: S65750; MUID:96049524; PMID:7578269
 A:Accession: S65750
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 57-69;97-110;111-129;204-213;216-244;255-271, 'X', 273-274;291-301 <PAC>
 C:Genetics:
 A:Gene: GDB:ELANH2; EI: P12
 A:Cross-references: GDB:132914; OMIM:130135
 A:Map position: 6p25-6p24.3
 C:Superfamily: antithrombin III
 C:Keywords: serine proteinase inhibitor

Query Match 74.7%; Score 62; DB 2; Length 379;
 Best Local Similarity 75.0%; Pred. No. 0.015;
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAT 16
 I::||| ||||| |||
 Db 322 SFVEVNEGTEAAAT 337

RESULT 5
 DYCH
 C:Species: Gallus gallus (chicken)
 C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 18-Jun-1999
 C:Accession: A01244
 R:Heilig, R.; Muraskowsky, R.; Kloefer, C.; Mandel, J.L.
 Nucleic Acids Res. 10, 4363-4382, 1982
 A:Title: The ovalbumin gene family: complete sequence and structure of the y gene.
 A:Reference number: A01244; MUID:83014329; PMID:7122240
 A:Accession: A01244
 A:Molecule type: DNA
 A:Residues: 1-388 <HEI>
 A:Cross-references: GB:J00922; GB:V00439; NID:9212899; PIDN:AAA68882.1; PID:9212900
 C:Genetics:
 A:Insertions: 56/3; 73/3; 116/3; 156/1; 203/3; 255/3
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; phosphoprotein; serine proteinase inhibitor
 F:4-121/disulfide bonds: #status predicted
 F:293/Binding site: carbonylate (Asn) (covalent) #status predicted
 F:345/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 72.3%; Score 60; DB 1; Length 388;
 Best Local Similarity 68.8%; Pred. No. 0.034;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 YIEVTEGTEATATG 17
 ::||| ||||| |||
 Db 332 FMEVNEGTEATGSGT 347

RESULT 6
 DXBHZ
 C:Species: Hordeum vulgare (barley)
 C:Date: 30-Jun-1987 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
 C:Accession: S13822; A01252
 R:Brandt, A.; Svendsen, I.; Hejgaard, J.
 Eur. J. Biochem. 194, 499-505, 1990
 A:Title: A plant serpin gene, structure, organization and expression of the gene enco
 A:Reference number: S13822; MUID:91099324; PMID:2269280
 A:Accession: S13822
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-399 <BRA>
 A:Cross-references: GB:X51726; NID:q19067; PIDN:CA36015.1; PID:q19068
 R:Hejgaard, J.; Rasmussen, S.K.; Brandt, A.; Svendsen, I.
 FEBS Lett. 180, 89-94, 1985
 A:Title: Sequence homology between barley endosperm protein Z and protease inhibitors

A:Reference number: A91332
A:Accession: A01252
A:Molecule type: mRNA
A:Residues: 220-399 <HEJ>
A:Cross-references: GB:X05902; NID:g19078; PIDN:CAA29331.1; PID:g19079
A:Note: most of the sequence has been confirmed by amino acid sequencing
R:Nielsen, G.; Johansen, H.; Jensen, J.; Hejgaard, J.
Barley Genet. Newsl. 13, 55-57, 1983
A:Title: Localization on barley chromosome 4 of genes coding for beta-amylase (Amy1) and
A:Reference number: A90095
A:Contents: annotation; map position
C:Comment: A major component of the endosperm albumin, this protein acts as a storage pr
ins of the antithrombin-III superfamily suggests that it also has an inhibitory function
C:Genetics:
A:Gene: Pz1
A:Map position: 4
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor; storage protein
F:357/inhibitory site: Met (unidentified proteinase) #status predicted

Query Match 72.38; Score 60; DB 1; Length 399;
Best Local Similarity 75.08; Pred. No. 0.035;
Matches 12; Conservative 2; Mismatches 0; Gaps 0;

Oy 1 SYIEVEEGTEATAT 16
Db 335 SFVEVNEEGTEAGNAT 350

RESULT 7
A57488
protease inhibitor Sp13 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 16-Jul-1999
C:Accession: A57488
R:Sun, J.; Rose, J.B.; Bird, P.
J. Biol. Chem. 270, 16089-16096, 1995
A:Title: Gene structure, chromosomal localization, and expression of the murine homology
A:Reference number: A57488; MUID:95332310; PMID:7608171
A:Accession: A57488
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <SUN>
A:Cross-references: GB:U25844; NID:g818902; PIDN:AAA79684.1; PID:g818903
C:Genetics:
A:Gene: Sp13
A:Map position: 13
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor
F:343/inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match 71.18; Score 59; DB 2; Length 378;
Best Local Similarity 68.88; Pred. No. 0.049;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SYIEVEEGTEATAT 16
Db 321 AFVEVNEEGTEAATAAT 336

RESULT 8
I38201
squamous cell carcinoma antigen 1 - human
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 26-May-2000
C:Accession: I38201; G01631; JT0966; JT0967
R:Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter, S.D.; Hui, S.
Proc. Natl. Acad. Sci. U.S.A. 92, 3147-3151, 1995
A:Title: A serine proteinase inhibitor locus at 18q21.3 contains a tandem duplication of
A:Reference number: I38200; MUID:95241462; PMID:7724531
A:Accession: I38201
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA

A:Residues: 1-390 <SCH1>
A:Cross-references: EMBL:U19568; GB:U19558; NID:g1172085
A:Accession: I38200
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-117 <SCH2>
A:Cross-references: EMBL:U19562; NID:g897835; PIDN:AAA6316.1; PID:g897844
R:Silverman, G.A.
submitted to the EMBL Data Library, January 1995
A:Reference number: G07968
A:Accession: G01631
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-350, 'G', 352-390 <STL>
A:Cross-references: EMBL:U19556; NID:g1276435; PID:g1052869
R:Sumitani, Y.; Kishi, F.; Sekiguchi, K.; Kato, H.
Biochem. Biophys. Res. Commun. 181, 51-58, 1991
A:Title: Squamous cell carcinoma antigen is a new member of the serine protease inhib
A:Reference number: JT0966; MUID:92068241; PMID:1958219
A:Accession: JT0966

A:Molecule type: mRNA
A:Residues: 1-350, 'G', 352-356, 'A', 358-390 <SUM1>
A:Cross-references: GB:566896; NID:g239551; PIDN:AAB20405.1; PID:g239552
A:Accession: JT0967
A:Molecule type: protein
A:Residues: 11-21, 231-237, 240-256, 303-325 <SUM2>
C:Comment: This antigen probably acts as a proteinase inhibitor to modulate the host
C:Genetics:
A:Gene: GDB:SCCA1; SCC
A:Cross-references: GDB:625364; OMIM:600517
A:Map position: 18q21.3-18q21.3
A:Residues: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3
C:Superfamily: antithrombin III
C:Keywords: cysteine proteinase inhibitor; glycoprotein
F:65,93,171,376/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:354/inhibitory site: Ser (cathepsin L) #status predicted

Query Match 71.18; Score 59; DB 2; Length 390;
Best Local Similarity 68.88; Pred. No. 0.05;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SYIEVEEGTEATAT 16
Db 332 AFVEVNEEGAEATAAT 347

RESULT 9
I38202
leupin precursor - human
N:Alternate names: proteinase inhibitor 11 (P11); squamous cell carcinoma antigen 2
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 26-May-2000
C:Accession: I38202; S66675; S57522
R:Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter, S.D.; Hui, S.
Proc. Natl. Acad. Sci. U.S.A. 92, 3147-3151, 1995
A:Title: A serine proteinase inhibitor locus at 18q21.3 contains a tandem duplication
A:Reference number: I38200; MUID:95241462; PMID:7724531
A:Accession: I38202
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-390 <SCH>
A:Cross-references: EMBL:U19576; GB:U19569; NID:g852466
R:Barnes, R.C.; Morrall, D.M.
FEBS Lett. 373, 61-65, 1995
A:Title: Identification of a novel human serpin gene; cloning sequencing and expressi
A:Reference number: S66675; MUID:96013887; PMID:7589435
A:Accession: S66675
A:Molecule type: mRNA
A:Residues: 7-351, 'V', 353-384 <BAR>
A:Cross-references: EMBL:X85015; NID:g887464; PIDN:CAA61420.1; PID:g887465
C:Genetics:
A:Gene: GDB:SCCA2; P11
A:Cross-references: GDB:636556; OMIM:600518

A:Map position: 18q21.3-18q21.3
 A:Intons: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; serine proteinase inhibitor
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-384/Product: leupin #status predicted <MAT>
 F:65,93,170,376/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:354/Inhibitory site: Leu (unidentified proteinase) #status predicted

Query Match 71.1%; Score 59; DB 2; Length 390;
 Best Local Similarity 68.8%; Pred. No. 0.05;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 STIEVTEGTEATAT 16
 DB 332 AFVEVTEGTEAATAAT 347
 ::::| | | | | | | | | |

RESULT 10
 A59273
 Proteinase inhibitor 8 - human
 N:Alternate names: cytoplasmic antiproteinase 2 (CAP-2)
 C:Species: Homo sapiens (man)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: A59273
 R: Sprecher, C.A.; Morgenstern, K.A.; Mathewes, S.; Dahlen, J.R.; Schrader, S.K.; Foster, J. Biol. Chem. 270, 29854-29861, 1995
 A:Title: Molecular cloning, expression, and partial characterization of two novel member
 A:Reference number: A59273; MUID:96102039; PMID:8530382
 A:Accession: A59273
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-374 <SPR>
 A:Cross-references: GB:I40377; NID:g1160926; PIDN:AC41939.1; PID:g1160927
 A:Experimental source: tissue type placenta; note (vector lambda gtl1); gene CAP2
 C:Genetics:
 A:Gene: GDB:PI8; CAP2
 A:Cross-references: GDB:599392; OMIM:601697
 A:Map position: 18q21.3-18q21.3
 C:Superfamily: antithrombin III
 C:Keywords: proteinase inhibitor
 F:339/Inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match 69.9%; Score 58; DB 2; Length 374;
 Best Local Similarity 73.3%; Pred. No. 0.071;
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 YIEVTEGTEATAT 16
 DB 318 FVEVTEGTEAATAAT 332
 ::::| | | | | | | | | |

RESULT 11
 B59273
 Proteinase inhibitor 9 - human
 N:Alternate names: cytoplasmic antiproteinase 3 (CAP-3)
 C:Species: Homo sapiens (man)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: B59273
 R: Sprecher, C.A.; Morgenstern, K.A.; Mathewes, S.; Dahlen, J.R.; Schrader, S.K.; Foster, J. Biol. Chem. 270, 29854-29861, 1995
 A:Title: Molecular cloning, expression, and partial characterization of two novel member
 A:Reference number: A59273; MUID:96102039; PMID:8530382
 A:Accession: B59273
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-376 <SPR>
 A:Cross-references: GB:I40378; NID:g1160928; PIDN:AC41940.1; PID:g1160929
 A:Experimental source: tissue type placenta; note (vector lambda gtl1); gene CAP2; tlassu
 C:Genetics:
 A:Gene: GDB:PI9; CAP3
 A:Cross-references: GDB:599393; OMIM:601799
 A:Map position: 6p25-6p25

C:Superfamily: antithrombin III
 C:Keywords: proteinase inhibitor
 F:340/Inhibitory site: Glu (unidentified proteinase) #status predicted

Query Match 69.9%; Score 58; DB 2; Length 376;
 Best Local Similarity 68.8%; Pred. No. 0.071;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 STIEVTEGTEATAT 16
 DB 318 SEVEVTEGTEAATAAS 333
 ::::| | | | | | | | | |

RESULT 12
 T06183
 serpin - barley
 C:Species: Hordeum vulgare (barley)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
 C:Accession: T06183
 R: Rasmussen, S.K.; Klausen, J.; Hejgaard, J.; Svenson, B.; Svendsen, I. Biochim. Biophys. Acta 1297, 127-130, 1996
 A:Title: Primary structure of the plant serpin B57 having the capacity of chymotryps
 A:Reference number: 215515; MUID:97075187; PMID:8917613
 A:Accession: T06183
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-397 <RAS>
 A:Cross-references: EMBL:X95277; NID:g1197576; PIDN:CAA64599.1; PID:g1197577
 A:Experimental source: cv. Boml, endosperm
 C:Superfamily: antithrombin III

Query Match 69.9%; Score 58; DB 2; Length 397;
 Best Local Similarity 64.7%; Pred. No. 0.075;
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 STIEVTEGTEATATG 17
 DB 336 FEVEDEGTEKAGATG 352
 ::::| | | | | | | | | |

RESULT 13
 AH1903
 hypothetical protein all0778 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AH1903
 R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH1903
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-374 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA872735.1; PID:g17130123; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all0778

Query Match 68.7%; Score 57; DB 2; Length 374;
 Best Local Similarity 85.7%; Pred. No. 0.1;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 IEVTEGTEATAT 16
 DB 315 IEVNEGTEASAT 328
 ::::| | | | | | | | | |

RESULT 14
 S29819
 serpin - barley

C:Species: Hordeum vulgare (barley)
 C:Date: 02-Dec-1993 #sequence_revision 12-Apr-1996 #text_change 16-Jul-1999
 C:Accession: S29819; S74166; S25625
 R:Rasmussen, S.K.
 Blochm. Biophys. Acta 1172, 151-154, 1993
 A:Title: A gene coding for a new plant serpin.
 A:Reference number: S29819; PMID:93176798; PMID:8439552
 A:Accession: S29819
 A:Molecule type: DNA
 A:Residues: 1-398 <RAS>
 A:Cross-references: EMBL:Z15116; NID:919070; PIDN:CAA78822.1; PID:919071
 A:Experimental source: strain Villa x Ulla
 R:Dahl, S.W.; Rasmussen, S.K.; Petersen, L.C.; Hejgaard, J.
 FEBS Lett. 394, 165-168, 1996
 A:Title: Inhibition of coagulation factors by recombinant barley serpin BSZx.
 A:Reference number: S74166; PMID:97000013; PMID:8843156
 A:Accession: S74166
 A:Molecule type: protein
 A:Residues: 357-361 <DAH>
 C:Genetics:
 A:Gene: pazx
 A:introns: 136/3
 C:Superfamily: antithrombin III

Query Match 68.7%; Score 57; DB 2; Length 398;
 Best Local Similarity 68.8%; Pred. No. 0.11;
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SYEVTEGTEATAT 16
 I::I I I I I I I I
 Db 334 SEVEVNEGTEAAART 349

RESULT 15
 A26423
 serine protease inhibitor 2.1 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 19-Nov-1988 #sequence_revision 12-Jul-1996 #text_change 12-Jul-1996
 C:Accession: A26423
 R:Hill, R.E.; Hastie, N.D.
 Nature 326, 96-99, 1987
 A:Title: Accelerated evolution in the reactive centre regions of serine protease inhibitor
 A:Reference number: A26423; PMID:87144617; PMID:3493437
 A:Accession: A26423
 A:Molecule type: mRNA
 A:Residues: 1-213 <HIL>
 A:Cross-references: GB:X13149
 A:Note: the authors translated the codon CTG for residue 1 as Met, CTG for residue 35 as
 11 as Met
 C:Superfamily: antithrombin III

Query Match 67.5%; Score 56; DB 2; Length 213;
 Best Local Similarity 73.3%; Pred. No. 0.084;
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 IEVTEGTEATATG 17
 ::I I I I I I I I I I
 Db 159 LDVAETGTETATATG 173

Search completed: July 11, 2003, 12:00:36
 Job time : 3.52827 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 1.62578 seconds

(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-4

Sequence: 1 SYIEVTEECTEATATG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	380	1	SPB7_HUMAN
2	73	88.0	378	1	PT16_BOVIN
3	69	83.1	392	1	SBI1_HUMAN
4	67	80.7	391	1	SBI3_HUMAN
5	62	74.7	376	1	PT16_HUMAN
6	62	74.7	379	1	ILEU_HUMAN
7	60	72.3	388	1	OVAY_CHICK
8	60	72.3	399	1	PRTE_HORVU
9	59	71.1	378	1	ILEU_PIG
10	59	71.1	380	1	PT16_MOUSE
11	59	71.1	390	1	SCC1_HUMAN
12	59	71.1	390	1	SCC2_HUMAN
13	58	69.9	374	1	SPB8_HUMAN
14	58	69.9	376	1	SPB9_HUMAN
15	57	68.7	410	1	NEUS_RAT
16	57	68.7	214	1	SI21_RAT
17	56	67.5	379	1	ILEU_HORSE
18	56	67.5	405	1	SBI2_HUMAN
19	56	67.5	410	1	NEUS_CHICK
20	56	67.5	410	1	NEUS_MOUSE
21	56	67.5	418	1	SI24_APOSY
22	55	66.3	410	1	NEUS_HUMAN
23	55	66.3	413	1	CP13_RAT
24	54	65.1	415	1	PA12_HUMAN
25	53	63.9	397	1	SBI0_HUMAN
26	52	62.7	222	1	SP1B_VACCC
27	52	62.7	341	1	SP12_COMPX
28	52	62.7	344	1	SP12_VARV
29	52	62.7	345	1	SP12_RABPU
30	52	62.7	345	1	SP12_VACCV
31	51	61.4	353	1	SP11_VACCC
32	51	61.4	353	1	SP11_VACCV
33	51	61.4	357	1	SP11_RABPU

34	51	61.4	397	1	GDN_MOUSE
35	51	61.4	398	1	GDN_HUMAN
36	51	61.4	418	1	COTR_MOUSE
37	51	61.4	418	1	CP16_RAT
38	51	61.4	433	1	ANT3_BOVIN
39	51	61.4	464	1	ANT3_HUMAN
40	51	61.4	465	1	ANT3_MOUSE
41	51	61.4	465	1	ANT3_SHEEP
42	50	60.2	355	1	SP11_COMPX
43	50	60.2	357	1	SP11_VARV
44	50	60.2	392	1	SERA_MOUSE
45	50	60.2	405	1	SP12_MOUSE

ALIGNMENTS

RESULT 1
ID SPB7_HUMAN STANDARD; PRT; 380 AA.
AC 075635;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Mgsin (TP55) (Serpin B7).
GN SERPINB7
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97326116; PubMed-9182567;
RA Tsujimoto M., Tsutouka N., Ishida N., Kurihara T., Iwasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino Y., Haruyama M., Mura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.;
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity."
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-98376492; PubMed-9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, mgsin, is a new serpin upregulated in
RT IgA nephropathy."
RL J. Clin. Invest. 102:828-836(1998).
CC -!- FUNCTION: Might influence the maturation of megakaryocytes via
CC proteases. Might influence the maturation of Lys-specific
CC its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: D88575; BAAJ1232.1; -
CC EMBL: AF027866; AAC64506.1; -
CC HSSP: P05619; 1HLE.
CC Genew: HGNC:13902; SERPINB7.
CC MTM: 603357; -
CC InterPro: IPR000215; Serpin;
CC Pfam: PF00079; serpin; 1.
CC SMART: SM00093; SERPIN; 1.
CC PROSITE: PS00284; SERPIN; 1.

KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 347 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFF605 CRC64;

Query Match 100.0%; Score 83; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATATG 17
 DB 325 SYIEVTEGTEATATG 341

RESULT 2

PT16_BOVIN STANDARD; PRT; 378 AA.
 AC 002739;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine proteinase inhibitor B-43.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 OX NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97165893; PubMed=9013786;
 RA Nakaya N., Nishibori M., Kawabata M., Saeki K.;
 RT "Cloning of a serine proteinase inhibitor from bovine brain:
 RT expression in the brain and characterization of its target
 RT proteinases.";
 RL Brain Res. Mol. Brain Res. 42:293-300(1996).
 CC PRESENT IN THE BRAIN OR EXTRAVASATED FROM THE BLOOD.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; D55670; BAA19875.1; -
 DR HSSP; P05120; IBY7.
 DR InterPro: IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 343 344 REACTIVE BOND (BY SIMILARITY).
 FT SEQUENCE 378 AA; 42560 MW; 664F499CCCE263A CRC64;

Query Match 88.0%; Score 73; DB 1; Length 378;
 Best Local Similarity 82.4%; Pred. No. 6.2e-05;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATATG 17
 DB 321 SFVEVTEGTEAATG 337

RESULT 3

SB11_HUMAN STANDARD; PRT; 392 AA.
 ID SB11_HUMAN
 AC Q96P15; Q96P14; Q96P13;
 DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serpin B11.
 GN SERPINB11.

OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS THR-148 AND THR-181.
 RA Silverman G.A., Cataltepe S.;
 RT "SERPINB11 is a novel ov-serpin with at least 3 alleles."
 RT Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF419953; AAL16056.1; -
 DR EMBL; AF419954; AAL16057.1; -
 DR EMBL; AF419955; AAL16058.1; -
 DR Genew; HGNC:14221; SERPINB11.
 DR InterPro: IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Polymorphism.
 FT ACT_SITE 357 358 REACTIVE BOND (BY SIMILARITY).
 FT ACT_SITE 148 148 M -> T (IN ALLELE B).
 FT VARIANT 181 181 /FTIG-Var_012472
 FT VARIANT 181 181 I -> T (IN ALLELE B AND ALLELE C).
 FT FTIG-Var_012473.
 FT SEQUENCE 392 AA; 44098 MW; 906F6DD412BCD756 CRC64;

Query Match 83.1%; Score 69; DB 1; Length 392;
 Best Local Similarity 76.5%; Pred. No. 0.00031;
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATATG 17
 DB 335 SYLDVSEGTETAAATG 351

RESULT 4

SB13_HUMAN STANDARD; PRT; 391 AA.
 ID SB13_HUMAN
 AC Q9U1V8; Q9UBW1; Q9UKG0; Q9HCX1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hurpin (Hacat UV-repressible serpin) (Protease inhibitor 13)
 DE (Headpin) (Serpin B13).
 GN SERPINB13 OR P113.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Skin;
 RX MEDLINE=99443898; PubMed=10512713;
 RA Abts H.F., Welts T., Mirmohammadsadeh A., Koehrer K., Michel G.,
 RA Ruzicka T.;
 RT "Cloning and characterization of hurpin (Protease inhibitor 13): a new
 RT skin specific, UV-repressible serine proteinase inhibitor of the
 RT ovalbumin serpin family.";
 RL J. Mol. Biol. 293:29-39(1999).
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE-9458661; PubMed-10527881;
 RA Spring P., Nakashima T., Frederick M., Henderson Y., Clayman G.;
 RT "Identification and cDNA cloning of headpin, a novel differentially
 RT expressed serpin that maps to chromosome 18q.";
 RL Biochem. Biophys. Res. Commun. 264:299-304(1999).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE-20461850; PubMed-11004515;
 RA Nakashima T., Pak S.C., Silverman G.A., Spring P.M., Frederick M.J.,
 RA Clayman G.L.;
 RT "Genomic cloning, mapping, structure and promoter analysis of HEADPIN,
 RT a serpin which is down-regulated in head and neck cancer cells.";
 RL Biochim. Biophys. Acta 1492:441-446(2000).
 RN (4)
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE-21213570; PubMed-11313015;
 RA Abts H.F., Wales T., Scheuring S., Scott F.L., Irving J.A., Michel G.,
 RA Bird P.I., Ruzicka T.;
 RT "Sequence, organization, chromosomal localization and alternative
 RT splicing of the human serine protease inhibitor gene hurpin (PI13),
 RT which is up-regulated in psoriasis.";
 RL DNA Cell Biol. 20:123-131(2001).
 CC -1- FUNCTION: May play a role in the proliferation or differentiation
 CC of keratinocytes.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2. are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SKIN-SPECIFIC.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ001696; CA04935.2; -;
 DR EMBL: AJ001697; CA04936.2; -;
 DR EMBL: AJ001698; CA04937.1; -;
 DR EMBL: AF169949; AAD5765.1; -;
 DR EMBL: AF169949; AAD5765.1; -;
 DR EMBL: AF278717; AAC72879.1; -;
 DR EMBL: AJ278717; CAC03569.1; -;
 DR HSSP: P05120; 1BY7.
 DR Genew: HGNC:8944; SERPINB13.
 DR MIM: 604445; -;
 DR InterPro: IPR000240; Maspin.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR PRINTS: PR00676; MASPIN.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 KM Serpin: Serine protease inhibitor; Alternative splicing.
 FT ACT_SITE 356 357 REACTIVE BOND (BY SIMILARITY).
 FT VARSPIC 206 257 MISSING (IN ISOFORM 2).
 FT CONFLICT 75 75 S -> N (IN REF. 4).
 FT CONFLICT 293 293 MISSING (IN REF. 1; CA04937).
 FT CONFLICT 297 297 G -> S (IN REF. 2).
 FT CONFLICT 297 297 E -> O (IN REF. 1; CA04937).
 SO SEQUENCE 391 AA; 44276 MW; 2CA88558D4BC2B09 CRC64;
 Query Match 80.7%; Score 67; DB 1; Length 391;
 Best Local Similarity 76.5%; Pred. No. 0.00067;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 5
 ID PT16_HUMAN STANDARD; PRT; 376 AA.
 AC P35237; Q96J44;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Placental thrombin inhibitor (cytoplasmic antiprotease) (CAP)
 DE (Protease inhibitor 6) (PI-6).
 GN SERPINB6 OR PI6 OR PTI.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE-94022386; PubMed-8415716;
 RA Coughlin P., Sun J., Cerruti L., Salem H.H., Bird P.;
 RT "Cloning and molecular characterization of a human intracellular
 RT serine proteinase inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9417-9421(1993).
 RN (2)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 67-73 AND 144-149.
 RC TISSUE-Placenta;
 RX MEDLINE-94183847; PubMed-8136380;
 RA Morgenstern K.A., Sprecher C.A., Holth L., Foster D., Grant F.J.,
 RA Ching A., Kistel W.;
 RT "Complementary DNA cloning and kinetic characterization of a novel
 RT intracellular serine proteinase inhibitor: mechanism of action with
 RT trypsin and factor Xa as model proteinases.";
 RL Biochemistry 33:3432-3441(1994).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Colon;
 RA Strausberg R.;
 RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Inhibits thrombin.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN SKELETAL MUSCLE. ALSO FOUND
 CC IN PLACENTA, CARDIAC MUSCLE, LUNG, LIVER, KIDNEY AND PANCREAS.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z22658; CA80373.1; -;
 DR EMBL: S69272; AAB30320.1; -;
 DR EMBL: BC001394; AAH01394.1; -;
 DR PIR: S35750; S35750.
 DR PIR: A46881; A46881.
 DR HSSP: P05120; 1BY7.
 DR Genew: HGNC:8950; SERPINB6.
 DR MIM: 173321; -;
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 KM Serpin: Serine protease inhibitor.
 FT ACT_SITE 341 342 REACTIVE BOND.
 FT CONFLICT 175 175 E -> G (IN REF. 1).
 FT CONFLICT 362 362 S -> R (IN REF. 1).
 SO SEQUENCE 376 AA; 42590 MW; 2B46A5540C608C6 CRC64;
 Query Match 74.7%; Score 62; DB 1; Length 376;
 Best Local Similarity 75.0%; Pred. No. 0.0045;
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEATAT 16
 ID ILEU_HUMAN STANDARD: PRT; 379 AA.
 AC P30740;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase inhibitor) (M/NEI) (EI).
 GN SERPINB1 OR ELANH2 OR P12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92302296; PubMed=1376927;
 RA Remold-O'Donnell E., Chin J., Alberts M.;
 RT "Sequence and molecular characterization of human monocyte/neutrophil elastase inhibitor.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:5635-5639(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9829265; PubMed=9630619;
 RA Zeng W., Silverman G.A., Remold-O'Donnell E.;
 RT "Structure and sequence of human M/NEI (monocyte/neutrophil elastase inhibitor), an Ov-serpin family gene.";
 RT Gene 213:179-187(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervix;
 RA Strusberg R.;
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 178-185; 204-210 AND 364-371.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E., Vandekeckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel protein database of normal human epidermal keratinocytes.";
 RT Electrophoresis 13:960-969(1992).
 CC -1- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASES ELASTASE, CATHEPSIN G AND PROTEINASE-3.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M93056; -, NOT_ANNOTATED_CDS.
 DR EMBL; AF053530; AAC31394.1; -.
 DR EMBL; BC009015; AAH09015.1; -.
 DR PIR; S27383; S27383.
 DR HSP; P05619; IHLF.
 DR Aarhus/Chent-2DPAGE; 4314; IEF.
 DR Genew; HGNC:3311; SERPINB1.
 DR MIM; 130135; -.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.

FT ACT SITE 344 345 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 379 AA; 42741 MW; BAAE08DFCBDC8D3 CRC64;
 Query Match
 Best Local Similarity 74.7%; Score 62; DB 1; Length 379;
 Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SYIEVTEEGTEATAT 16
 ID ILEU_HUMAN STANDARD: PRT; 388 AA.
 AC P01014;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1993 (Rel. 25, last annotation update)
 DE Gene Y protein (Ovalbumin-related).
 GN Y.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 NCBI_TaxID=9031;
 (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83014329; PubMed=7122240;
 RA Helling R., Muraskowsky R., Kloepper C., Mandel J.L.;
 RT "The ovalbumin gene family: complete sequence and structure of the Y gene.";
 RT Nucleic Acids Res. 10:4363-4382(1982).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; J00922; AAA68882.1; -.
 DR PIR; A01244; DYCH.
 DR HSP; P01012; IOVA.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin.
 FT ACT SITE 353 354 REACTIVE BOND HOMOLOGY.
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 388 AA; 43772 MW; 2AE17BF715D7E461 CRC64;
 Query Match
 Best Local Similarity 72.3%; Score 60; DB 1; Length 388;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 2 YIEVTEEGTEATATG 17
 ID PRZ_HORVU STANDARD: PRT; 399 AA.
 AC P06293;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1991 (Rel. 20, last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein 2 (24) (Major endosperm albumin).
 GN PAZ1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Hordeum.
 NCBI_TaxID=4513;
 RN
 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-CV, Carlsberg II; TISSUE=Grain;
 RA MEDLINE=9109324; PubMed=2269280;
 RA Brandt A., Svendsen I., Hejgaard J.;
 RT "A plant serpin gene. Structure, organization and expression of the
 RT gene encoding barley protein 24.";
 RL Eur. J. Biochem. 194:499-505(1990).
 RN
 [2]
 RP SEQUENCE OF 220-399 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-CV, Carlsberg II; TISSUE=Grain;
 RA Hejgaard J., Rasmussen S.K., Brandt A., Svendsen I.;
 RT "Sequence homology between barley endosperm protein 2 and protease
 RT inhibitors of the alpha-1-antitrypsin family.";
 RL FEBS Lett. 180:89-94(1985).
 RN
 [3]
 RP SEQUENCE OF 360-384 FROM N.A.
 RA Rasmussen S.K., Hopp H.E., Brandt A., Svendsen I., Hejgaard J.;
 RT "A cDNA clone for protein 2, a major barley endosperm albumin.";
 RL Carlsberg Res. Commun. 49:385-390(1984).
 CC
 -1- FUNCTION: A MAJOR COMPONENT OF THE ENDOSPERM ALBUMIN, THIS PROTEIN
 CC ACTS AS A STORAGE PROTEIN DURING GRAIN FILLING, CONTRIBUTING A
 CC SUBSTANTIAL PART OF THE GRAIN'S LYSINE. MAY HAVE AN INHIBITORY
 CC FUNCTION DURING FILLING OR GERMINATION.
 CC
 -1- TISSUE SPECIFICITY: IS ACCUMULATED AND STORED IN THE ENDOSPERM,
 CC WHERE IT EXISTS IN A FREE AND A BOUND FORM.
 CC
 -1- DEVELOPMENTAL STAGE: SYNTHESIZED 10-25 DAYS AFTER FERTILIZATION
 CC (DEVELOPING ENDOSPERM).
 CC
 -1- INDUCTION: ITS EXPRESSION IS REGULATED BY THE "HIGH LYSINE"
 CC ALLELES LYS1 AND LYS3A.
 CC
 -1- MISCELLANEOUS: THERE SEEM TO BE TWO 2 PROTEINS: 24 (FROM
 CC CHROMOSOME 4) AND 27 (FROM CHROMOSOME 7).
 CC
 -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

 DR EMBL: X51726; CAA36015.1; -;
 DR EMBL: X05902; CAA29331.1; -;
 DR EMBL: M35065; AAA32971.1; -;
 DR PIR: A01352; DMBH2.
 DR PIR: S13822; S13822.
 DR HSSP: P01008; IATH.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Plant defense; Serpin; Storage protein; Multigene family.
 FT DOMAIN 36 56
 FT SIGNAL FOR TARGETING PROTEIN 24 INTO
 FT THE ER LUMEN (POTENTIAL).
 FT REACTIVE BOND (POTENTIAL).
 FT ACT_SITE 357 357
 FT SEQUENCE 399 AA; 43376 MW; 023BBE52475BAA56 CRC64;
 QY Query Match 72.3%; Score 60; DB 1; Length 399;
 DB Best Local Similarity 75.0%; Pred. No. 0.011;
 DB Matches 12; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 9
 ID ILEU_PIG STANDARD; PRT; 378 AA.
 AC P80229;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte elastase inhibitor (LEI) (Leukocyte neutral proteinase
 DE inhibitor) (UNP).
 GN SERPINB1 OR ELANH2.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 RN
 [1]
 RP SEQUENCE.
 RC TISSUE=Leukocyte;
 RX MEDLINE=94039085; PubMed=7901009;
 RA Teschauer W.F., Mentle R., Sommerhoff C.P.;
 RT "Primary structure of a porcine leukocyte serpin.";
 RL Eur. J. Biochem. 217:519-526(1993).
 CC
 -1- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASE
 CC ELASTASE, CATEPSIN G AND PROTEINASE-3 (By similarity).
 CC
 -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC
 -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC
 -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 DR HSSP: P05619; IHLE.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; serine protease inhibitor.
 FT MOD_RES 1
 FT MOD_RES 1
 FT DISULFID 80 80
 FT ACT_SITE 343 344
 FT SEQUENCE 378 AA; 42512 MW; 3CA7C65F11E508CA CRC64;
 QY Query Match 71.1%; Score 59; DB 1; Length 378;
 DB Best Local Similarity 68.8%; Pred. No. 0.015;
 DB Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 10
 ID PT16_MOUSE STANDARD; PRT; 378 AA.
 AC Q60834;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Placental thrombin inhibitor (Protease inhibitor 6) (PI-6).
 GN SERPINB6 OR SPI3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=95332310; PubMed=7608171;
 RA Sun J., Rose J.B., Bird P.;
 RT "Gene structure, chromosomal localization, and expression of the
 RT murine homologue of human proteinase inhibitor 6 (PI-6) suggests
 RT divergence of PI-6 from the ovalbumin serpins.";
 RL J. Biol. Chem. 270:16089-16096(1995).
 RN
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;

RA Strausberg R.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Inhibits thrombin.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U25844; AAA79684.1; -
 DR EMBL: BC006766; AAH06766.1; -
 DR HSSP: P05120; 1BY7.
 DR MGD: MGI:103123; Serpinb6.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KM Serpin; Serine protease inhibitor.
 FT ACT_SITE 343 344 42599 MW; 4B0F5E1A030BDF6 CRC64;
 SQ SEQUENCE 378 AA; 42599 MW; 4B0F5E1A030BDF6 CRC64;
 Query Match 71.1%; Score 59; DB 1; Length 378;
 Best Local Similarity 68.8%; Pred. No. 0.015;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SYTEVTEGETEATAT 16
 DB 321 AFVEVTEGEAAAT 336
 RESULT 11
 SCC1_HUMAN STANDARD; PRT; 390 AA.
 ID SCC1_HUMAN
 AC P29508; Q96J21;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Squamous cell carcinoma antigen 1 (SCCA-1) (Protein T4-A).
 GN SERPINB3 OR SCCA1 OR SCCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE-92068241; PubMed-1958219;
 RA Sumitani Y., Kishi F., Sekiguchi K., Kato H.;
 RT "Squamous cell carcinoma antigen is a new member of the serine
 RT protease inhibitors.";
 RL Biochem. Biophys. Res. Commun. 181:51-58(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95241462; PubMed-7724531;
 RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.;
 RA Treter S.D., Hui S.M., Silverman G.A.;
 RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem
 RT duplication of the human squamous cell carcinoma antigen gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
 CC IMMUNE RESPONSE AGAINST TUMOR CELLS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.
 CC -1- DEVELOPMENTAL STAGE: ITS EXPRESSION IS CLOSELY RELATED TO CELLULAR

CC DIFFERENTIATION IN BOTH NORMAL AND MALIGNANT SQUAMOUS CELLS.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S66896; AAB20405.1; -
 DR EMBL: U19556; AAA97552.1; -
 DR EMBL: U19568; AAA86317.1; -
 DR EMBL: U19559; AAA86317.1; JOINED.
 DR EMBL: U19562; AAA86317.1; JOINED.
 DR EMBL: U19565; AAA86317.1; JOINED.
 DR EMBL: U19567; AAA86317.1; JOINED.
 DR EMBL: U19562; AAA86316.1; -
 DR EMBL: U19559; AAA86316.1; JOINED.
 DR EMBL: U19560; AAA86316.1; JOINED.
 DR EMBL: BC005224; AAH05224.1; -
 DR PIR: J70966; J70966.
 DR HSSP: P01008; IATR.
 DR GeneW: HGNC:10569; SERPINB3.
 DR MIM: 600517; -
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KM Serpin; Serine protease inhibitor.
 FT ACT_SITE 354 355
 FT CONFLICT 357 357 REACTIVE BOND.
 FT FT T -> A (IN REF. 1).
 SQ SEQUENCE 390 AA; 44564 MW; E5F27P986C752CFA CRC64;
 Query Match 71.1%; Score 59; DB 1; Length 390;
 Best Local Similarity 68.8%; Pred. No. 0.015;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SYTEVTEGETEATAT 16
 DB 332 AFVEVTEGEAAAT 347
 RESULT 12
 SCC2_HUMAN STANDARD; PRT; 390 AA.
 ID SCC2_HUMAN
 AC P48594;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Squamous cell carcinoma antigen 2 (SCCA-2) (Lepin).
 GN SERPINB4 OR SCCA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95241462; PubMed-7724531;
 RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.;
 RA Treter S.D., Hui S.M., Silverman G.A.;
 RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem
 RT duplication of the human squamous cell carcinoma antigen gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96013887; PubMed-7589435;
 RA Barnes R.C., Morrall D.M.;
 RT "Identification of a novel human serpin gene: cloning sequencing and
 RT expression of Lepin.";
 RL FEBS Lett. 373:61-65(1995).

[3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=21167379; PubMed=11267667;
 RA Hamada K., Shimomiyu H., Asano Y., Kihana T., Iwamoto M., Hanakawa Y.,
 RA Hashimoto K., Hirose S., Ito M.;
 RT "Molecular cloning of human squamous cell carcinoma antigen 1 gene and
 RT characterization of its promoter.";
 RL Blochlm. Biophys. Acta 1518:124-131(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA Strauberg R.;
 RU Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
 CC IMMUNE RESPONSE AGAINST TUMOR CELLS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U19576; AAA92602.1; -;
 DR EMBL: U19570; AAA92602.1; JOINED.
 DR EMBL: U19571; AAA92602.1; JOINED.
 DR EMBL: U19572; AAA92602.1; JOINED.
 DR EMBL: U19574; AAA92602.1; JOINED.
 DR EMBL: U19575; AAA92602.1; JOINED.
 DR EMBL: U19557; AAA97553.1; -;
 DR EMBL: X89015; CA61420.1; -;
 DR EMBL: AB035089; BAB21525.1; -;
 DR EMBL: BC017401; AAH17401.1; -;
 DR HSSP: P01008; IATH.
 DR Genev. HGNC:10570; SERPINB4.
 DR MIM: 600518; -;
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin. 1.
 DR SMART: SM00093; SERPIN. 1.
 DR PROSITE: PS00284; SERPIN. 1.
 KW Serpin: Serine protease inhibitor.
 FT ACT_SITE 354 355
 FT SEQUENCE 390 AA; 44854 MW; 04E213CD89258705 CRC64;
 SO
 Query Match 71.1%; Score 59; DB 1; Length 390;
 Best Local Similarity 68.8%; Pred. NO. 0.015;
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RC TISSUE=Placenta;
 RX MEDLINE=96102039; PubMed=8530382;
 RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,
 RA Schrader S.K., Foster D.C., Kistiel W.;
 RT "Molecular cloning, expression, and partial characterization of two
 RT novel members of the ovalbumin family of serine proteinase
 RT inhibitors.";
 RL J. Biol. Chem. 270:29854-29861(1995).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: LA0377; AAC41939.1; -;
 DR HSSP: P05120; 1BY7
 DR Genev. HGNC:8952; SERPINB8.
 DR MIM: 601697; -;
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin. 1.
 DR SMART: SM00093; SERPIN. 1.
 DR PROSITE: PS00284; SERPIN. 1.
 KW Serpin: Serine protease inhibitor.
 FT ACT_SITE 339 340
 FT SEQUENCE 374 AA; 42786 MW; E855A033519AC60E CRC64;
 SO
 Query Match 69.9%; Score 58; DB 1; Length 374;
 Best Local Similarity 73.3%; Pred. NO. 0.022;
 Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 2 YIEVTEGETEATAT 16
 Db 318 FVEVTEGETEATAT 332

RESULT 14
 SPB9_HUMAN STANDARD: PRT; 376 AA.
 AC P50453;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytoplasmic antiprotease 3 (CAP3) (CAP-3) (Protease inhibitor 9)
 DE (Serpin B9).
 GN SERPINB9 OR P19.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96102039; PubMed=8530382;
 RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,
 RA Schrader S.K., Foster D.C., Kistiel W.;
 RT "Molecular cloning, expression, and partial characterization of two
 RT novel members of the ovalbumin family of serine proteinase
 RT inhibitors.";
 RL J. Biol. Chem. 270:29854-29861(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta; and Bone marrow;
 RX MEDLINE=97066975; PubMed=8910377;
 RA Sun J., Bird C.H., Sutton V., McDonald L., Coughlin P.B., Jong T.A.,
 RA Trepant J.A., Bird P.I.;
 RT "A cytosolic granzyme B inhibitor related to the viral apoptotic
 RT regulator cytokine response modifier A is present in cytotoxic
 RT lymphocytes.";

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 7.42204 Seconds
(Without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-4
Perfect score: 83
Sequence: 1 SYIEVTEEGTEATATAC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteriap:*
17: sp._archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	86.7	334	12	Q9DHG4 yaba-like d
2	69	83.1	380	11	Q9D695 mus musculu
3	69	83.1	388	11	Q9CQV3 mus musculu
4	69	83.1	388	11	Q9A122 mus musculu
5	68	81.9	379	11	Q9D7S8 mus musculu
6	68	81.9	379	11	Q9D1S4 mus musculu
7	68	81.9	387	11	Q9D1O5 mus musculu
8	68	81.9	388	11	Q9Z2G2 mus musculu
9	66	79.5	379	5	Q9NF76 anopheles g
10	66	79.5	379	5	Q9NF76 anopheles g
11	66	79.5	380	5	Q9NF76 anopheles g
12	66	79.5	380	5	Q9NF76 anopheles g
13	66	79.5	380	5	Q9NF76 anopheles g
14	66	79.5	392	5	Q9U1I5 drosophila
15	66	79.5	418	5	Q8T0M5 drosophila
16	66	79.5	424	5	Q917G6 drosophila

17	66	79.5	431	5	Q9V991	Q9V991 drosophila
18	65	78.3	380	11	Q920J5	Q920J5 ratu
19	62	74.7	177	4	Q9BSM3	Q9BSM3 homo sapien
20	62	74.7	280	10	Q9XET7	Q9XET7 avena fatua
21	62	74.7	374	5	Q9V3L3	Q9V3L3 drosophila
22	62	74.7	377	11	Q08804	Q08804 mus musculu
23	62	74.7	382	11	Q8VHP7	Q8VHP7 mus musculu
24	62	74.7	411	10	Q94DW6	Q94DW6 oryza sativ
25	61	73.5	374	5	Q9NGS0	Q9NGS0 bombyx mori
26	60	72.3	359	11	Q9DGA7	Q9DGA7 mus musculu
27	60	72.3	382	5	Q8WSX9	Q8WSX9 anopheles g
28	60	72.3	395	5	Q8WSX7	Q8WSX7 anopheles g
29	60	72.3	400	10	Q40076	Q40076 hordium vul
30	59	71.1	338	4	Q9BYR8	Q9BYR8 homo sapien
31	59	71.1	369	4	Q9BYR7	Q9BYR7 homo sapien
32	59	71.1	377	5	Q9GP13	Q9GP13 Ixodes ricin
33	59	71.1	378	5	Q8WQX1	Q8WQX1 rhipicephal
34	59	71.1	379	11	Q9D0S8	Q9D0S8 mus musculu
35	59	71.1	385	11	Q8VHQ0	Q8VHQ0 mus musculu
36	58	69.9	374	11	Q08800	Q08800 mus musculu
37	58	69.9	380	5	Q8WQX0	Q8WQX0 rhipicephal
38	58	69.9	386	11	Q9D1E7	Q9D1E7 mus musculu
39	58	69.9	397	10	Q43492	Q43492 hordium vul
40	57	68.7	225	6	Q9GMA9	Q9GMA9 sus scrofa
41	57	68.7	373	5	Q9NH65	Q9NH65 trichinella
42	57	68.7	374	16	Q8VXR8	Q8VXR8 anabaena sp
43	57	68.7	398	10	Q40066	Q40066 hordium vul
44	57	68.7	415	6	Q9GMA6	Q9GMA6 sus scrofa
45	57	68.7	459	11	Q9WTT1	Q9WTT1 cavia porce

ALIGNMENTS

RESULT 1
ID Q9DHG4 PRELIMINARY; PRT; 334 AA.
AC Q9DHG4:
DT 01-MAR-2001 (TREMUREL. 16, Created)
DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)
DE 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE 149R protein.
GN 149R.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-2176366; PubMed-11277691;
RA Lee H.J.; Essani K.; Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Theiss (2000), Sir William Dunn School of Pathology, University of.
RL Thesis (2000), Sir William Dunn School of Pathology, University of.
CC EMBL: AJ293568; CAC21387.1; -.
DR HSSP; P05120; 1877.
DR InterPro; IPR00215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 334 AA; 38190 MW; D7FD004191B3C19F CRC64;
Query Match 86.7%; Score 72; DB 12; Length 334;
Best Local Similarity 93.8%; Pred. No. 0.00075;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATAT 16
 |||||
 Db 277 SYIEVTEGTEAATAAT 292

RESULT 2
 ID 09D695 PRELIMINARY; PRT; 380 AA.
 AC 09D695;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 4631416M05R1K protein (megsin).
 GN SERPINB7 OR 4631416M05R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SKIN;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Noridone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 Hayashizaki Y.,
 RA Hayshtzaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21368006; PubMed=11473647;
 RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
 Yagi M., Nagano N., Inagi R., Kurokawa K.,
 RT "Cloning of rodent mesgin revealed its up-regulation in
 RT mesangioliproliferative nephritis."
 RL Kidney Int. 60:641-652(2001).
 DR EMBL; AK014524; BAB29410.1; -
 DR EMBL; AF105328; AAL16768.1; -
 DR HSSP; P05121; 1ATC.
 DR MGD; MGI:2151053; Serpinb7.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; UNKNOWN.1.
 SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CF4 CRC64;

Query Match 83.1%; Score 69; DB 11; Length 380;
 Best Local Similarity 87.5%; Pred. No. 0.0027;
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATAT 16
 |||||
 Db 325 SYIEVSECTEATAT 340

ID 09CQV3 PRELIMINARY; PRT; 388 AA.
 AC 09CQV3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 2310046M08R1K protein.
 GN 2310046M08R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Noridone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 Hayashizaki Y.,
 RA Hayshtzaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK009855; BAB26546.1; -
 DR EMBL; AK009003; BAB26017.1; -
 DR HSSP; P05619; 1HLE.
 DR MGD; MGI:1914207; 2310046M08R1K.
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 SQ SEQUENCE 388 AA; 43482 MW; 533D802BF58F09AE CRC64;

Query Match 83.1%; Score 69; DB 11; Length 388;
 Best Local Similarity 76.5%; Pred. No. 0.0028;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
 ||:| |||||
 Db 332 SYVDNECTEATAATG 348

RESULT 4
 ID 091212 PRELIMINARY; PRT; 388 AA.
 AC 091212;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE RIKEN cDNA 2310046M08 gene.
 GN 2310046M08R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.,
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 CC -i- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL; BC010313; AAH10313.1; -
 DR MGD; MGI:1914207; 2310046M08R1K.

DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin. 1.
KW Serpin.
SQ SEQUENCE 388 AA; 43544 MW; AF4BB62F8084E6FA CRC64;

Query Match 83.1%; Score 69; DB 11; Length 388;
Best Local Similarity 76.5%; Pred. No. 0.0028;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEAATG 17
DB 332 SFVEVNEEGTEAAATG 348

RESULT 5

AC 09D758 PRELIMINARY; PRT; 379 AA.
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE 1190005M04R1k protein.
GN SERPINB1A OR 1190005M04R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AK008914; BAB25964.1; -;
DR HSSP: P05619; 1HLE.
DR MGD: MGI:1913472; Serpinb1a.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin. 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 379 AA; 42580 MW; 193728585077AB4B CRC64;

Query Match 81.9%; Score 68; DB 11; Length 379;
Best Local Similarity 76.5%; Pred. No. 0.004;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEAATG 17
DB 322 SFVEVNEEGTEAAATG 338

RESULT 6

09D154

ID 09D154 PRELIMINARY; PRT; 379 AA.

AC 09D154;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE 1190005M04R1k protein (RIKEN cDNA 1190005M04 gene) (E1A).
GN SERPINB1A OR 1190005M04R1K OR SERPINB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AK003930; BAB23079.1; -;
DR EMBL: BC011140; AAH11140.1; -;
DR EMBL: AF426024; AAL57486.1; -;
DR HSSP: P05619; 1HLE.
DR MGD: MGI:1913472; Serpinb1a.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin. 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 379 AA; 42575 MW; F492F1EB88BE80B1 CRC64;

Query Match 81.9%; Score 68; DB 11; Length 379;
Best Local Similarity 76.5%; Pred. No. 0.004;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEAATG 17
DB 322 SFVEVNEEGTEAAATG 338

RESULT 7

09D105

RA	Silverman G. A.;
RT	"A murine ortholog of the human serpin SCC2 maps to chromosome 1 and inhibits chymotrypsin-like-serine proteinases.";
RL	Genomics 54:297-306(1998).
CC	-I - SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR	EMBL; AF063937; AAC95432.1; .
DR	HSSP; P05120; Ibv7.
DR	MED; MG1:1277952; Soc2.
DR	InterPro; IPR000215; Serpin.
DR	Pfam; PF00079; serpin; 1.
DR	SMART; SM00093; SERPIN; 1.
DR	PROSITE; PS00284; SERPIN; 1.
KW	Serpin.
KW	SEQUENCE 388 AA; 44573 MW; D77AFAL23EB9CD CRC64;
QY	1 SYIEVTEEGTEATPATG 17 :::
Db	331 SFVEVNEGTGAATAAGT 347
QY	Query Match 81.9%; Score 68; DB 11; Length 388; Best Local Similarity 76.5%; Pred. No. 0.0041; Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
RESULT 9	
Q9NFT6	PRELIMINARY; PTR; 379 AA.
ID Q9NFT6	
AC Q9NFT6	
DT 01-OCT-2000 (TREMBLrel_15, Created)	
DT 01-OCT-2000 (TREMBLrel_15, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel_21, Last annotation update)	
DE Putative serine protease inhibitor.	
GN Sp1A.	
OS Anopheles gambiae (African malaria mosquito).	
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;	
OC Anopheles.	
OX NCBI_TaxID=7165;	
ON [1]	
RN RP	SEQUENCE FROM N.A.
RC STRAIN=AA R/R;	
RA Danieli A., Kafatos F.C.;	
RT "Serine protease inhibitors of Anopheles gambiae.";	
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AJ271352; CAB69784.1; .	
DR HSSP; P01008; IATH.	
DR InterPro; IPR000215; Serpin.	
DR Pfam; PF00079; serpin; 1.	
DR SMART; SM00093; SERPIN; 1.	
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.	
KW Protease.	
KW SEQUENCE 379 AA; 42157 MW; 3DE2DFOA07BCB959 CRC64;	
QY	Query Match 79.5%; Score 66; DB 5; Length 379; Best Local Similarity 76.5%; Pred. No. 0.0087; Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db	1 SYIEVTEEGTEATPATG 17 :::
Db	323 AFIEVNEGTGAATAAGT 339
QY	Query Match 81.9%; Score 68; DB 11; Length 388; Best Local Similarity 76.5%; Pred. No. 0.0041; Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
RESULT 10	
Q8MSYO	PRELIMINARY; PTR; 379 AA.
ID Q8MSYO	
AC Q8MSYO	
DT 01-MAR-2002 (TREMBLrel_20, Created)	
DT 01-MAR-2002 (TREMBLrel_20, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel_21, Last annotation update)	
DE Serpin.	
GN SP12IF.	
OS Anopheles gambiae (African malaria mosquito).	
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RN SEQUENCE FROM N.A.
RA Danieli A., Kafatos T.G., Kafatos F.C.;
RT "Isolation and characterization of four alternatively spliced serpin
RT variants in the mosquito Anopheles gambiae."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ2420785; CAD12781.1; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SO SEQUENCE 379 AA; 42161 MW; 3F29524C13BC468 CRC64;

Query Match 79.5%; Score 66; DB 5; Length 379;
Best Local Similarity 76.5%; Pred. No. 0.0087;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SYIEVEEGTEAATG 17
Db 323 AFIEVNEGTEAATG 339
:::|||||

RESULT 11
ID Q9NFT5 PRELIMINARY; PRT; 380 AA.
AC Q9NFT5;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative serine protease inhibitor.
GN SP1B.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RN SEQUENCE FROM N.A.
RA Danieli A., Kafatos T.G.;
RT "Serine protease inhibitors of Anopheles gambiae";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ271353; CAB69785.1; -
DR HSP: P01008; IATV.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
KW Protease.
SO SEQUENCE 380 AA; 42577 MW; EF4C5A581BEC276 CRC64;

Query Match 79.5%; Score 66; DB 5; Length 380;
Best Local Similarity 76.5%; Pred. No. 0.0087;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SYIEVEEGTEAATG 17
Db 323 AFIEVNEGTEAATG 339
:::|||||

RESULT 12
ID Q8WSX8 PRELIMINARY; PRT; 380 AA.
AC Q8WSX8;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Serpin.
GN SP12F.
OS Anopheles gambiae (African malaria mosquito).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RN SEQUENCE FROM N.A.
RA Danieli A., Loukeris T.G., Kafatos F.C.;
RT "Isolation and characterization of four alternatively spliced serpin
RT variants in the mosquito Anopheles gambiae."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ420785; CAD12783.1; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SO SEQUENCE 380 AA; 42582 MW; C3FE91D55DF24508 CRC64;

Query Match 79.5%; Score 66; DB 5; Length 380;
Best Local Similarity 76.5%; Pred. No. 0.0087;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SYIEVEEGTEAATG 17
Db 323 AFIEVNEGTEAATG 339
:::|||||

RESULT 13
ID Q917G5 PRELIMINARY; PRT; 392 AA.
AC Q917G5;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Sp4 protein.
GN SP4 OR CG9453.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Daveport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Padle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Mu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Ruidin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003790; AAC22209.1; -
 DR HSSP: P01008; IATH.
 DR FLYBase: FBgn0028985; sp4.
 DR InterPro: IPR000886; ER_target.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
 SQ SEQUENCE 392 AA; 44127 MW; 5D221529C5E0822E CRC64;

Query Match 79.5%; Score 66; DB 5; Length 392;
 Best Local Similarity 76.5%; Pred. No. 0.009;

Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
 DB 319 AFIEVNEGTAAATG 335

RESULT 14

090115

ID 090115 PRELIMINARY; PRT; 392 AA.

AC 090115; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Serine protease inhibitor (Serpin-4).

GN SP4 OR CG9453.

OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RA Han J.H., Zhang H., Min G.S., Hashimoto C.,

RT "Identification of functional *Drosophila* serpins."

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

DR EMBL: A0251747; CAB63099.1; -

DR HSSP: P01008; IATH.

DR FLYBase: FBgn0028985; sp4.

DR InterPro: IPR000886; ER_target.

DR InterPro: IPR000215; Serpin.

DR Pfam: PF00079; serpin.1.

DR SMART: SM00093; SERPIN.1.

DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.

KW Protease; Serpin.

SQ SEQUENCE 392 AA; 44133 MW; 0A01CF14D3B85B9F CRC64;

Query Match 79.5%; Score 66; DB 5; Length 392;
 Best Local Similarity 76.5%; Pred. No. 0.009;

Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
 DB 319 AFIEVNEGTAAATG 335

RESULT 15

08T0M5

ID 08T0M5 PRELIMINARY; PRT; 418 AA.

AC 08T0M5; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE GH21896p.
 GN SP4.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Broksstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Paclob J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY069173; AAL39318.1; -
 SQ SEQUENCE 418 AA; 46961 MW; D956A2C04339B5E0 CRC64;

Query Match 79.5%; Score 66; DB 5; Length 418;
 Best Local Similarity 76.5%; Pred. No. 0.0096;

Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
 DB 358 AFIEVNEGTAAATG 374

Search completed: July 11, 2003, 11:58:57
 Job time: 10.422 secs

GenCore version 5.1.6
Copyright (c) 1993
using sw model
16424 Second
Arguments)
million cell upda

908470

Scoring tab
atch 1008
first 45 summaries

SUMMARIES

Description	

human megakaryocyt
human megasin prote
human megasin prote
human megakaryocyt
human megasin assoc
Human GTPase assoc
Human human diagno
Novel C. amalonaticus ma
Lettuce resistance
Lettuce pest resis

ALIGNMENTS

Lettuce resistance
Lettuce pest resis
Drosophila melanog
C albicans apoptos
C albicans pylon
Helicobacter pylor
Pseudomonas aerugi
P. furiosus endonuc
Novel human protel
Novel human protein
Novel humaprotein
propiolin secretum
Chlamydia pneumonia
Human ORFX protein
pyroclitium occult
Drosophila melanog
Drosop secreted pro
Human
Aspergillus niger
Drosophila melanog
E. blattae 1,3-p
E. blattae fastidiosa
Xylella polypleptide
Human protein sequ
Human CDPK (clone)
Rice yeast DNA pol y
P. abduyi DNA mult
P. occultum occult
Pyroclitium occult
Human myosin IIA
Novel human diapor
Diplolepis tremblac
H. pylori ORF 028P
H. pylori ORF
Human protein sequ
Arbidolopsis thallic

RESULT 1	
AA848379	standard; protein; 380 AA.
ID	AA848379
XX	PAR48379;
XX	16-AUG-1994 (first entry)
XX	Human megakaryocyte differentiation factor; MDF; thrombopoietin.
XX	Human megakaryocyte differentiation factor; thrombocytopenia; platelet.
XX	Human megakaryocytic differentiation factor; cancer chemotherapy.
XX	haematopoietic stimulation; bone marrow transplantation;
KM	bone marrow transplantation; cancer chemotherapy.
XX	Homo sapiens.
OS	EP583884-A.
XX	23-FEB-1994.
PN	93EP-0305654.
XX	19-JUL-1993;
PD	92JP-0212305.
XX	17-JUL-1992;
PF	93JP-0067339.
XX	04-MAR-1993;
PR	(SUNR) SUNTORY LTD.
XX	(TSUJ/) TSUJIMOTO M.
PA	(TSUJ/) TSUJIMOTO M.
PA	Iwasa F, Kurihara T, Miura K, Nakazato H;
XX	Ishida N, Tsuruoka N, Yamaguchi N, Yamachi K;
XX	Tsuji moto M.
PI	WPI; 1994-058782/08.
XX	WPI; 1994-058782/08.
DR	N-PSDB; AA05670.

us-10-091-442-5.1rag

Fri Jul 11 15:11:04 2003

XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX Claim 7: Page 30-32; 47pp; English.

XX Human MDF can be isolated from a culture of human epidermoid
XX carcinoma A431 cells in protein-free medium. The MDF stimulates
CC differentiation of megakaryocytes from myeloid cells in the presence
CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC for treatment of diseases involving a decrease in platelet number
CC (esp. thrombocytopenia) such as occurs in bone marrow
CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by
CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.

XX Sequence 380 AA: 100.0%; Score 44; DB 15; Length 380;
XX Query Match Similarity 100.0%; Pred. No. 1.1; Indels 0; Gaps 0;
XX Best Local Similarity 100.0%; Mismatches 0; Conservative

Matches 9; Conservative
OY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 2
ID AAY08254 standard: Protein: 380 AA.
XX AAY08254:
XX 14-JUL-1999 (first entry)
XX Human megalin protein.
XX Megalin: mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
XX human; rat; murine.
XX no sapiens.

XX 652-A1.
XX DR 9 N-98WO-JP04269.
XX PT Megalin pI-0275302.
XX PS Claim 1: Page 62.
XX CC This invention describes
XX CC proteins from human, rat and
XX CC specifically in mesangial cells
XX CC useful for the treatment and diagnosis
XX CC cells, such as Iga nephropathy.
XX Sequence 380 AA:
XX Query Match
XX Best Local Similarity 100.0%; Score 44; DB 20; Length 380;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0

OY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 3
ID AAB24142 standard: Protein: 380 AA.
XX AAB24142:
XX 30-JAN-2001 (first entry)

XX Human megalin protein sequence SEQ ID NO:2.
XX Megalin: mesangium-predominant gene; serpin regulated; nephropathy;
XX Iga; immunoglobulin A; detection; renal function; renal disorder;
XX diagnosis; biological sample; blood; urine.

XX Homo sapiens.
XX WO200057189-A1.
XX 28-SEP-2000.
XX 17-MAR-2000; 2000WO-JP01646.
XX 19-MAR-1999; 99JP-0075305.
XX 28-OCT-1999; 99JP-0306623.
XX (KURO/) KUROKAWA K.
XX (FUSO) FUSO PHARM IND LTD.
XX (MIYA/) MIYATA T.
XX Miyata T.
XX WPI: 2000-611642/58.
XX N-PSB; AAB99294.
XX DR 9 N-PSB; AAB99294.
XX PT Evaluating renal function comprises assaying megalin protein in
XX biological sample
XX Example 2: Page 66-69; 93pp; Japanese.

XX The present invention describes a method for evaluating renal function.
XX CC The method comprises assaying megalin protein in biological sample. Also
XX CC described are: (1) use of a anti-megalin protein antibody for diagnosing;
XX CC renal function; and (2) a kit for detecting megalin protein particles;
XX CC (a) anti-megalin protein antibody attached to the particles; and
XX CC (b) direct or indirect fixing for the antibody to the particles;
XX CC (c) a magnet. The process is useful for evaluating renal function and
XX CC diagnosing renal disorders by assaying megalin protein in biological
XX CC samples (preferably urine or blood). The process represents the human megalin
XX CC protein, which is given in the exemplification of the present invention
XX Sequence 380 AA:
XX Query Match
XX Best Local Similarity 100.0%; Score 44; DB 21; Length 380;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0

OY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 4
ID AAB83075 standard: Protein: 380 AA.
XX AAB83075:
XX AAB83075:
XX AAB83075:

```

XX 10-JUL-2001 (first entry)
XX
XX Human mesgin protein.
DE
XX Human; mesgin; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening.
XX
XX Homo sapiens.
OS
XX WO200124628-A1.
PN
XX 12-APR-2001.
PD
XX
XX 06-OCT-2000: 2000WO-JP06988.
PF
XX
XX 06-OCT-1999: 99JP-0285736.
PR
XX (KURO/) KUROKAWA K.
PA (MIVA/) MIYATA T.
XX
XX Miyata T;
PI
XX WPI: 2001-300136/31.
DR N-PSDB; AAF82438.
XX
XX Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments
PS Example 4; Page 44-46; 62pp; Japanese.
XX
XX The present sequence is human mesgin. The human mesgin coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.
XX
XX Sequence 380 AA:
SQ
XX
XX Query Match 100.0%; Score 44; DB 22; Length 380;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QYRALGLK 9
DB 289 QYRALGLK 297

```

```

XX 23-FEB-1994.
XX
XX 19-JUL-1993: 93EP-0305654.
XX
XX 17-JUL-1992: 92JP-0212305.
PR 04-MAR-1993: 93JP-0067339.
XX
XX (SUNR ) SUNTORY LTD.
PA (TSUJ/) TSUJIMOTO M.
XX
XX Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
XX
XX WPI: 1994-058782/08.
DR
XX
XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
XX Claim 1; Page 19; 47pp; English.
PS
XX Human MDF (see AAR48379) can be isolated from a culture of human
CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
CC stimulates differentiation of megakaryocytes from myeloid cells
CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
CC making it useful for treatment of diseases involving a decrease
CC in platelet number (esp. thrombocytopenia) such as occurs in bone
CC marrow transplantation and in chemotherapy. MDF has mol.wt.
CC 55-57KD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
CC contains an amino acid sequence comprising at least one of the
CC sequences AAR57107-R57115.
XX
XX Sequence 9 AA:
SQ
XX
XX Query Match 88.6%; Score 39; DB 15; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 YLRALGLK 9
DB 2 YLRALGLK 9

```

```

RESULT 5
AAR57111
ID AAR57111 standard; peptide; 9 AA.
XX
XX AAR57111;
AC
XX 16-AUG-1994 (first entry)
DT
XX
XX Human megakaryocyte differentiation factor peptide 5.
DE
XX
XX Human megakaryocyte differentiation factor; MDF; thrombopoietin;
KW haematopoietic stimulating factor; thrombocytopoena; platelet;
KW bone marrow transplantation; cancer chemotherapy.
XX
XX Homo sapiens.
OS
XX
XX key Location/Qualifiers
FH Misc-difference 1 /note= "not determined"
FT
XX EP583884-A.
PN

```

```

RESULT 6
AAY99661
ID AAY99661 standard; Protein; 694 AA.
XX
XX AAY99661;
AC
XX 03-NOV-2000 (first entry)
DT
XX
XX Human GTPase associated protein-12.
DE
XX
XX Guanine nucleotide binding protein; GTP-binding protein; G-protein;
KW GTPase; GTPase associated protein; GTPAP; cell proliferation;
KW autoimmune; inflammatory; immune system disorder; cancer; AIDS;
KW acquired immune deficiency syndrome; asthma; atherosclerosis;
KW arthritis; systemic lupus erythematosus; psoriasis; human.
XX
XX Homo sapiens.
OS
XX
XX key Location/Qualifiers
FH Domain 10..24
FT /label= Beta_transducin
FT Modified-site 57
FT /note= "Potential phosphorylation site"
FT Modified-site 67
FT /note= "Potential phosphorylation site"
FT Domain 96..110
FT /label= Beta_transducin
FT Modified-site 99
FT /note= "Potential phosphorylation site"

```

FT	Modified-site	150	/note=	"Potential phosphorylation site"
FT	Modified-site	156	/note=	"Potential phosphorylation site"
FT	Modified-site	209	/note=	"Potential phosphorylation site"
FT	Modified-site	285	/note=	"Potential phosphorylation site"
FT	Modified-site	331	/note=	"Potential phosphorylation site"
FT	Modified-site	346	/note=	"Potential phosphorylation site"
FT	Modified-site	360	/note=	"Potential phosphorylation site"
FT	Modified-site	388	/note=	"Potential phosphorylation site"
FT	Modified-site	416	/note=	"Potential phosphorylation site"
FT	Modified-site	430	/note=	"Potential phosphorylation site"
FT	Modified-site	467	/note=	"Potential phosphorylation site"
FT	Modified-site	477	/note=	"Potential phosphorylation site"
FT	Modified-site	500	/note=	"Potential phosphorylation site"
FT	Modified-site	522	/note=	"Potential phosphorylation site"
FT	Modified-site	650	/note=	"Potential phosphorylation site"
FT	Modified-site	684	/note=	"Potential phosphorylation site"
FT	Modified-site	688	/note=	"Potential phosphorylation site"
FT	Modified-site		/note=	"Potential phosphorylation site"
PN	WO200031263-A2.			
XX	02-JUN-2000.			
XX	23-NOV-1999;	99WO-US28013.		
XX	23-NOV-1998;	98US-0109592.		
PR	04-FEB-1999;	99US-0118610.		
PR	06-APR-1999;	99US-0127990.		
XX	(INCY-)	INCYTE PHARM INC.		
XX	Hillman JL, Tang YT, Bandman O, Lal P, Yue H, Lu DAM, Baughn MR;			
PI	Yang J, Azimzal Y;			
XX	WPI: 2000-400073/34.			
DR	N-PSDB: AAA49182.			
XX	Human GTPase associated proteins, polynucleotides, and antibodies,			
PT	useful for diagnosing, preventing and treating various diseases such as			
PT	atherosclerosis, cancer, acquired immune deficiency syndrome (AIDS),			
PT	asthma, and autoimmune diseases -			
XX	Claim 1; Page 94-96; 144pp: English.			
PS	Human cDNA libraries from various tissues were screened for GTPase			
CC	associated proteins (GTPAP). The present sequence is human			
CC	GTPAP-12 protein. This sequence was derived from a cDNA library of the			
CC	prostate tumour tissue removed from a 65 year old male.			
CC	This protein is expressed in reproductive, haematopoietic/immune and			
CC	gastrointestinal tissue. The GTPAP proteins may be used to define			
CC	agonists and antagonists of GTPAP activity and to generate antibodies			
CC	to GTPAP. This means the GTPAP proteins may be useful for treatment or			
CC	prevention of diseases associated with GTPAP such as cell proliferation			
CC	disorders, autoimmune disorders, inflammatory disorders, immune system			
CC	disorders, cancer, AIDS, asthma, atherosclerosis, arthritis, systemic			
CC	lupus erythematosus and psoriasis.			
XX				

[illegible]

Query Match 77.3%; Score 34; DB 22; Length 213;
 Best Local Similarity 87.5%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGL 8
 DB 142 OYLRALGL 149

RESULT 8

AA06788
 ID AA06788 standard; Protein; 250 AA.

AC AA06788;

DT 22-JUN-1999 (first entry)

DE C. amalonaticus maleate isomerase.

KM Maleate isomerase; recombinant; P. fluorescens; C. amalonaticus;
 KW L-aspartic acid.

OS Citrobacter amalonaticus.

PN EP908520-A2.

PD 14-APR-1999.

PF 08-OCT-1998; 98EP-0308213.

PR 08-OCT-1997; 97JP-0276261.

PA (JAPC) NIPPON SHOKUBAI CO LTD.

PI Komatsuzaki S, Mukoyama M, Yasuda S;

DR WPI; 1999-217071/19.

DR N-PSDB; AAX32470.

PT New maleate isomerase useful for production of L-aspartic acid

PS Claim 6; Page 11-12; 21pp; English.

XX This represents a C. amalonaticus maleate isomerase. Host cells

CC transformed with a vector comprising the DNA are used for the recombinant

CC expression of the protein. The maleate isomerases of the invention

CC isolated from P. fluorescens and C. amalonaticus are useful for

CC producing L-aspartic acid.

XX

SO Sequence 250 AA;

OY 1 OYLRALGL 9

DB 228 OYLRALGL 236

AA064160
 ID AA064160 standard; Protein; 1365 AA.

AC AA064160;

DT 19-JUL-1999 (first entry)

DE Lettuce resistance protein RLGA.

KM Resistance gene; RLGA; RG2A; lettuce; disease resistance;

KW pest resistance; virus; fungus; protozoan; bacterium; nematode;
 crop protection; transgenic plant.

XX Lactuca sativa.

OS Lactuca sativa.

XX Key Location/Qualifiers

FT Misc-difference 413 /note- "encoded by RAC"

FT Misc-difference 434 /note- "encoded by NTG"

FT Misc-difference 1057 /note- "encoded by NAG"

FT Misc-difference 1086 /note- "encoded by NAG"

XX WO9830083-A1.

PN 16-JUL-1998.

PF 09-JAN-1998; 98WO-US00615.

PR 10-JAN-1997; 97US-0781734.

PA (REGC) UNIV CALIFORNIA.

PI Meyers B, Michelmore RW, Shen K;

DR WPI; 1998-398692/34.

DR N-PSDB; AAV44199.

PT New resistance gene nucleic acids - useful to produce disease

PT resistant plants, e.g. lettuce, Lactuca sativa, and to detect

PT resistance genes, e.g. to fingerprint cultivars

XX Claim 37; Page 77; 183pp; English.

XX The present sequence represents lettuce RLGA resistance protein,

CC as deduced from the RLGA gene (see AAV44199) isolated from lettuce

CC cultivar Diana genomic DNA. RG (or RLGA) families RG1-RG5 and RG7

CC have been identified, each RG family being defined as a group of

CC polypeptide sequences that have at least 60% amino acid sequence

CC identity. Individual members of an RG family typically map to the

CC same genomic locus. RG nucleic acid sequences (see AAV4418-257) can

CC be used to confer resistance in plants to a variety of pests,

CC including viruses, fungi, nematodes, insects and bacteria.

CC Sequences from within the RG genes can be used to fingerprint

CC cultivars or genoplasm for the presence of desired resistance

CC genes. Antibodies to RG proteins (see AAV44150-98) are useful for

CC detecting the type and amount of RG protein expressed in a plant

XX sample.

SO Sequence 1365 AA;

OY 1 OYLRALGL 9

DB 31 OYLRALGL 39

AA095473
 ID AA095473 standard; Protein; 1366 AA.

AC AA095473;

DT 02-JUL-2002 (first entry)

DE Lettuce pest resistance protein RLGA #1.

KM Plant; lettuce; resistance gene; plant downy mildew; RG; transgenic.

XX Lactuca sativa.

XX US6350933-B1.
 XX 26-FEB-2002.
 XX 09-JAN-1998; 98US-0004838.
 XX 10-JAN-1997; 97US-0781734.
 XX (REGC) UNIV CALIFORNIA.
 XX Michelmore RW, Shen KA, Meyers BC;
 XX WPI: 2002-314681/35.
 XX N-PSDB; ABR67796.
 XX Isolated nucleic acid molecule, RG2B (resistance gene 2B)
 XX polynucleotide, encodes polypeptide useful for conferring resistance to
 XX plant downy mildew -
 XX
 XX Disclosure: Column 85-92; 209pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule comprises an
 XX RG2B (resistance gene 2B) polynucleotide encoding a polypeptide
 XX conferring resistance to plant downy mildew, hybridises to a fully
 XX defined sequence of 15062 base pairs as given in the specification, under
 XX a wash with 0.2 x SSC (standard sodium citrate) at 65plusoc. Also
 XX included is a transgenic plant comprising a heterologous expression
 XX cassette containing a promoter operably linked to an RG2B polynucleotide
 XX encoding a polypeptide conferring resistance to plant downy mildew and
 XX hybridises to a fully defined sequence of 1323 amino acids as
 XX given in the specification, under a wash with 0.2 x SSC at 65plusoc.
 XX Also disclosed are other pest resistance genes and their protein
 XX products. The isolated nucleic acid molecule is used to confer resistance
 XX to downy mildew in a plant. This comprises introducing into the plant
 XX an expression cassette comprising a promoter operably linked to an RG2B
 XX polynucleotide. The present sequence is a lettuce resistance
 XX protein sequence (or protein consensus sequence).
 XX
 XX Sequence 1366 AA;
 XX
 XX Query Match 75.0%; Score 33; DB 23; Length 1366;
 XX Best Local Similarity 55.6%; Pred. No. 6.5e+02;
 XX Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX 1 QYRALGLK 9
 XX 11:1:1:1
 XX 31 QYRMGMGIX 39
 XX
 XX RESULT 11
 XX AAM64176
 XX ID AAM64176 standard; Protein; 1890 AA.
 XX
 XX AAM64176;
 XX
 XX 19-JUL-1999 (first entry)
 XX
 XX Lettuce resistance protein RG2A.
 XX
 XX Resistance gene; RG2A; lettuce; disease resistance;
 XX pest resistance; virus; fungus; protozoan; bacterium; nematode;
 XX crop protection; transgenic plant.
 XX
 XX Lactuca sativa.
 XX
 XX Key Location/Qualifiers
 XX FT Misc-difference 1497
 XX FT /note= "encoded by TAA"
 XX FT Misc-difference 1884
 XX FT /note= "encoded by TAA"
 XX
 XX W09830083-A1.

XX 16-JUL-1998.
 XX 09-JAN-1998; 98WO-US000615.
 XX 10-JAN-1997; 97US-0781734.
 XX (REGC) UNIV CALIFORNIA.
 XX Meyers B, Michelmore RW, Shen K;
 XX WPI: 1998-398692/34.
 XX N-PSDB; AAV44231.
 XX New resistance gene nucleic acids - useful to produce disease
 XX resistant plants, e.g. lettuce, Lactuca sativa, and to detect
 XX resistance genes, e.g. to fingerprint cultivars
 XX
 XX Claim 37; Page 110; 183pp; English.
 XX
 XX The present sequence represents lettuce RG2A resistance protein,
 XX as deduced from the RG2A gene (see AAV44231) isolated from lettuce
 XX cultivar Diana genomic DNA. RG (or RUG) families RG1-RG5 and RG7
 XX have been identified, each RG family being defined as a group of
 XX polypeptide sequences that have at least 60% amino acid sequence
 XX identity. Individual members of an RG family typically map to the
 XX same genomic locus. RG nucleic acid sequences (see AAV44188-257) can
 XX be used to confer resistance in plants to a variety of pests,
 XX including viruses, fungi, nematodes, insects and bacteria.
 XX Sequences from within the RG genes can be used to fingerprint
 XX cultivars or germplasm for the presence of desired resistance
 XX genes. Antibodies to RG proteins (see AAM64150-98) are useful for
 XX detecting the type and amount of RG protein expressed in a plant
 XX sample.
 XX
 XX Sequence 1890 AA;
 XX
 XX Query Match 75.0%; Score 33; DB 19; Length 1890;
 XX Best Local Similarity 55.6%; Pred. No. 9e+02;
 XX Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX 1 QYRALGLK 9
 XX 11:1:1:1
 XX 31 QYRMGMGIX 39
 XX
 XX RESULT 12
 XX AAU95491
 XX ID AAU95491 standard; Protein; 1890 AA.
 XX
 XX AAU95491;
 XX
 XX 02-JUL-2002 (first entry)
 XX
 XX Lettuce pest resistance protein RG2A #3.
 XX
 XX Plant; lettuce; resistance gene; plant downy mildew; RG; transgenic.
 XX
 XX Lactuca sativa.
 XX
 XX US6350933-B1.
 XX 26-FEB-2002.
 XX 09-JAN-1998; 98US-0004838.
 XX 10-JAN-1997; 97US-0781734.
 XX (REGC) UNIV CALIFORNIA.
 XX Michelmore RW, Shen KA, Meyers BC;
 XX WPI: 2002-314681/35.

DR N-PSDB; ABK67843.
XX Isolated nucleic acid molecule, RG2B (resistance gene 2B)
PT polynucleotide, encodes polypeptide useful for conferring resistance to
PT plant downy mildew
XX
PS Disclosure; Column 209-220: 209pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising an
CC RG2B (resistance gene 2B) polynucleotide encoding a polypeptide
CC conferring resistance to plant downy mildew, hybridises to a fully
CC defined sequence of 15062 base pairs as given in the specification, under
CC a wash with 0.2 x SSC (standard sodium citrate) at 65plusOC. Also
CC included is a transgenic plant comprising a heterologous expression
CC cassette containing a promoter operably linked to an RG2B polynucleotide
CC encoding a polypeptide conferring resistance to plant downy mildew and
CC hybridises to a fully defined sequence of 1323 amino acids as
CC given in the specification, under a wash with 0.2 x SSC at 65plusOC.
CC Also disclosed are other pest resistance genes and their protein
CC products. The isolated nucleic acid molecule is used to confer resistance
CC to downy mildew in a plant. This comprises introducing into the plant
CC an expression cassette comprising a promoter operably linked to an RG2B
CC polynucleotide. The present sequence is a lettuce resistance
CC protein sequence (or protein consensus sequence).
XX
SQ Sequence 1890 AA:

Query Match 75.0%; Score 33; DB 23; Length 1890;
Best Local Similarity 55.6%; Pred. No. 9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
||:|:|:|
Db 31 QYRREKGIK 39

RESULT 13
ABK66978
ID ABK66978 standard; Protein: 194 AA.
XX
AC ABB66978;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 27726.
XX
KW Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, LI PWD, Myers EW;
XX
DR WPI: 2001-6556860/75.
XX
DR N-PSDB; ABL11081.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 27726; 21pp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 194 AA:

Query Match 72.7%; Score 32; DB 22; Length 194;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
||:|:|:|
Db 50 QYCRSLGLQ 58

RESULT 14
AAG70876
ID AAG70876 standard; Protein: 281 AA.
XX
AC AAG70876;
XX
DT 27-JUL-2001 (first entry)
XX
DE C albicans apoptosis associated protein #56.
XX
KW Yeast; fungus; apoptosis; infection; proliferative disease;
XX vaccine; autoimmune disease; ischaemia; neurodegeneration.
XX
OS Candida albicans.
XX
PN WO200102550-A2.
XX
PD 11-JAN-2001.
XX
PE 03-JUL-2000; 2000WO-BE00077.
XX
PR 01-JUL-1999; 99EP-0870141.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
XX Nelissen BJM, Reekmans RJ;
XX
DR WPI: 2001-367042/38.
XX
DR N-PSDB; AAH29912.
XX
XX Yeast and fungal nucleic acids encoding proteins involved in a pathway
PT leading to programmed cell death, useful for treating proliferative
PT disorders; yeast and fungal infections, or for preventing apoptosis in
PT certain diseases -
XX
PS Claim 24; Fig 2; 218pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of apoptosis associated proteins from the yeast Saccharomyces
CC cerevisiae and the fungus Candida albicans. These can be used to identify
CC treatments for fungal and yeast infections, for proliferative diseases
CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
CC and neurodegeneration. The present sequence is one of the C. albicans
CC proteins of the invention.
XX
SQ Sequence 281 AA:

Query Match 72.7%; Score 32; DB 22; Length 281;
Best Local Similarity 75.0%; Pred. No. 2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
||| ||:|
Db 95 YLRALGLK 102

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYRALGLK 9
:|| ||||:
Db 313 EYRALGLQ 321

Search completed: July 11, 2003, 11:54:21
Job time : 6.16424 secs

RESULT 15
AAU35870
ID AAU35870 standard; Protein; 327 AA.

XX AAU35870;

XX 14-FEB-2002 (first entry)

DE Helicobacter pylori cellular proliferation protein #183.

XX Antisense; prokaryotic cellular proliferation protein;

KM antibiotic; antibacterial; drug design.

XX Helicobacter pylori.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX MPI: 2001-611495/70.

DR N-PSDB; AAS53729.

XX

PT New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

PS Example 3; Seq ID No 11463; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 327 AA:

Query Match 72.7%; Score 32; DB 22; Length 327;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 1.60915 Seconds
(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-5
Perfect score: 44
Sequence: 1 OYRALGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/p/odacata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/p/odacata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/p/odacata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/p/odacata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/p/odacata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/p/odacata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	2	US-08-611-977-5
2	44	100.0	380	2	US-08-472-659-34
3	44	100.0	380	2	US-08-474-661-34
4	44	100.0	380	2	US-08-611-977-34
5	39	88.6	9	2	US-08-472-659-5
6	39	88.6	9	2	US-08-474-661-5
7	33	75.0	250	4	US-09-167-117-2
8	33	75.0	1366	4	US-09-004-838-22
9	33	75.0	1890	4	US-09-004-838-88
10	31	70.5	387	4	US-09-314-847A-4
11	31	70.5	387	4	US-09-570-778A-2
12	31	70.5	387	4	US-09-570-778A-7
13	31	70.5	387	4	US-09-570-778A-8
14	31	70.5	387	4	US-09-570-778A-9
15	31	70.5	623	4	US-09-347-801-4
16	31	70.5	623	4	US-09-347-801-4
17	31	70.5	803	1	US-08-062-368-2
18	31	70.5	803	1	US-08-062-368-4
19	31	70.5	2548	4	US-09-172-422-1
20	31	68.2	223	4	US-08-896-410-4
21	30	68.2	310	4	US-09-071-035-412
22	30	68.2	327	4	US-08-896-410-38
23	30	68.2	347	4	US-09-071-035-410
24	30	68.2	387	4	US-08-968-563-37
25	30	68.2	387	4	US-08-969-583A-37
26	30	68.2	387	4	US-09-570-778A-10
27	30	68.2	387	4	US-09-570-778A-11

28	30	68.2	393	4	US-09-134-001C-5594	Sequence 5594, App
29	30	68.2	512	4	US-09-134-078-26	Sequence 26, Appl
30	30	68.2	533	4	US-09-360-197-10	Sequence 10, Appl
31	30	68.2	1327	4	US-08-196-387-2	Sequence 2, Appl1
32	29	65.9	10	1	US-08-300-864-60	Sequence 60, Appl
33	29	65.9	10	3	US-08-931-645-60	Sequence 60, Appl
34	29	65.9	10	5	PCT-US94-01258-60	Sequence 60, Appl
35	29	65.9	10	5	PCT-US95-11235-60	Sequence 60, Appl
36	29	65.9	214	4	US-09-228-986-129	Sequence 129, App
37	29	65.9	223	4	US-08-928-941D-22	Sequence 22, Appl
38	29	65.9	223	4	US-09-280-590A-22	Sequence 22, Appl
39	29	65.9	315	4	US-09-439-261-15	Sequence 15, Appl
40	29	65.9	315	4	US-09-227-613-16	Sequence 16, Appl
41	29	65.9	349	2	US-07-857-224B-95	Sequence 95, Appl
42	29	65.9	372	4	US-08-928-941D-16	Sequence 16, Appl
43	29	65.9	372	4	US-08-280-590A-16	Sequence 16, Appl
44	29	65.9	441	1	US-09-287-937C-11	Sequence 11, Appl
45	29	65.9	446	2	US-08-833-610-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-611-977-5
Sequence 5, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
INVENTOR: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5972886uo
APPLICANT: NAKAZOTO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: KISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
CORRESPONDENCE ADDRESSES:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-5

Query Match 100.0%; Score 44; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 1 QYLRALGK 9

RESULT 2
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 44; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 3
US-08-474-661-34

Sequence 34, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 44; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 4

US-08-611-977-34
: Sequence 34, Application US/08611977
: Patent No. 5972886
: GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masaaki
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSUROUOKA, No. 5972886uo
: APPLICANT: NAKAZATO, Hiroshi
: APPLICANT: MIURA, Kenju
: APPLICANT: KURIHARA, Tatsuya
: APPLICANT: YAMAGUCHI, Kozo
: APPLICANT: YAMAGUCHI, Kozo
: TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Burns, Doane, Swecker & Mathis
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/611,977
: FILING DATE: 06-MAR-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION NUMBER:
: APPLICATION NUMBER: US 08/091,028
: FILING DATE: 14-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 4-212305
: FILING DATE: 17-JUL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-067339
: FILING DATE: 04-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: McGowan, Malcolm K.
: REGISTRATION NUMBER: 39,300
: REFERENCE/DOCKET NUMBER: 001560-204
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 34:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 380 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-611-977-34

Query Match 100.0%; Score 44; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YLRALGLK 9
Db 289 YLRALGLK 297

RESULT 5
US-08-472-659-5
: Sequence 5, Application US/08472659
: Patent No. 5831030
: GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masaaki
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSUROUOKA, No. 5831030uo
: APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030unhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-472-659-5

Query Match 88.6%; Score 39; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YLRALGLK 9
Db 2 YLRALGLK 9

RESULT 6
US-08-474-661-5
: Sequence 5, Application US/08474661
: Patent No. 5874253
: GENERAL INFORMATION:
: APPLICANT: TSUJIMOTO, Masaaki
: APPLICANT: IWASA, Fuyuki
: APPLICANT: TSUROUOKA, No. 5874253uo
: APPLICANT: NAKAZATO, Hiroshi
: APPLICANT: MIURA, Kenju
: APPLICANT: KURIHARA, Tatsuya
: APPLICANT: ISHIDA, No. 5874253unhiro
: APPLICANT: KURIHARA, Tatsuya
: APPLICANT: YAMAGUCHI, Kozo
: APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: George Mason Bldg., Washington & Prince Sts.
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,661
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 5-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REA, TERESA STANER
;; REGISTRATION NUMBER: 30,427
;; REFERENCE/DOCKET NUMBER: 001560-204
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-6620
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-474-661-5

Query Match 88.6%; Score 39; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
Db 2 YLRALGLK 9

RESULT 7
US-09-167-717-2
; Sequence 2, Application US/09167717
; Patent No. 6133014
; GENERAL INFORMATION:
; APPLICANT: MUKOYAMA, MASAHARU
; APPLICANT: YASUDA, SHINZO
; APPLICANT: KOMATSUZAKI, SATOMI
; TITLE OF INVENTION: MALEATE ISOMERASE GENE
; FILE REFERENCE: PH-555US
; CURRENT APPLICATION NUMBER: US/09/167,717
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: JP 276261/1997
; EARLIER FILING DATE: 1997-10-08
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Citrobacter amalonaticus
US-09-167-717-2

Query Match 75.0%; Score 33; DB 4; Length 250;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGLK 9
Db 228 OMKALGLK 236

RESULT 8
US-09-004-838-22
; Sequence 22, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelmore, Richard W.
; APPLICANT: Shen, Kathy
; TITLE OF INVENTION: Procedures and Materials for
; TITLE OF INVENTION: Confering Pest Resistance in Plants
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,838
; FILING DATE: 09-JAN-1998
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/761,734
; FILING DATE: 10-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 023070-078810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1366 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..1366
; OTHER INFORMATION: /note="RLG2a amino acids".
US-09-004-838-22

Query Match 75.0%; Score 33; DB 4; Length 1366;
Best Local Similarity 55.6%; Pred. No. 86;
Matches 5; Conservative .3; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGLK 9
Db 31 OYREMGIK 39

RESULT 9
US-09-004-838-88
; Sequence 88, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelmore, Richard W.
; APPLICANT: Shen, Kathy
; APPLICANT: Meyers, Blake

TITLE OF INVENTION: Procedures and Materials for
NUMBER OF INVENTION: Confering Pest Resistance In Plants
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,838
FILING DATE: 09-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/781,734
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Elmhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 023070-078810US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 88:
SEQUENCE CHARACTERISTICS:
LENGTH: 1890 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY:
LOCATION: 1..1890
OTHER INFORMATION: /note- "RC2A deduced sequence"
US-09-004-838-88
Query Match
Best Local Similarity 75.0%; Score 33; DB 4; Length 1890;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 31 OYKREMGK 39
RESULT 10
US-09-314-847A-4
Sequence 4, Application US/09314847A
Patent No. 6365410
GENERAL INFORMATION:
APPLICANT: Schellenberger, Volker
APPLICANT: Liu, Amy D.
APPLICANT: Sellionova, Olga V.
TITLE OF INVENTION: Directed Evolution of Microorganisms
FILE REFERENCE: GC560
CURRENT APPLICATION NUMBER: US/09/314,847A
CURRENT FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 387
TYPE: PRT
ORGANISM: E. blatte
US-09-314-847A-4
Query Match 70.5%; Score 31; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 342 OHLRELGVK 350
RESULT 11
US-09-570-778A-2
Sequence 2, Application US/09570778A
Patent No. 6468773
GENERAL INFORMATION:
APPLICANT: Trimbur, Donald E.
APPLICANT: Whited, Gregory M.
APPLICANT: Sellionova, Olga V.
TITLE OF INVENTION: Mutant 1,3-propanediol Dehydrogenase
FILE REFERENCE: GC580-2
CURRENT APPLICATION NUMBER: US/09/570,778A
CURRENT FILING DATE: 2000-05-14
PRIOR APPLICATION NUMBER: US 60/134,868
PRIOR FILING DATE: 1999-05-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 387
TYPE: PRT
ORGANISM: Escherichia blataae
US-09-570-778A-2
Query Match 70.5%; Score 31; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 342 OHLRELGVK 350
RESULT 12
US-09-570-778A-7
Sequence 7, Application US/09570778A
Patent No. 6468773
GENERAL INFORMATION:
APPLICANT: Trimbur, Donald E.
APPLICANT: Whited, Gregory M.
APPLICANT: Sellionova, Olga V.
TITLE OF INVENTION: Mutant 1,3-propanediol Dehydrogenase
FILE REFERENCE: GC580-2
CURRENT APPLICATION NUMBER: US/09/570,778A
CURRENT FILING DATE: 2000-05-14
PRIOR APPLICATION NUMBER: US 60/134,868
PRIOR FILING DATE: 1999-05-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 387
TYPE: PRT
ORGANISM: Escherichia blataae
US-09-570-778A-7
Query Match 70.5%; Score 31; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 342 OHLRELGVK 350
RESULT 13
US-09-570-778A-8
Sequence 8, Application US/09570778A
Patent No. 6468773

```
; GENERAL INFORMATION:
; APPLICANT: Trimbur, Donald E.
; APPLICANT: Whited, Gregory M.
; APPLICANT: Selfonova, Olga V.
; TITLE OF INVENTION: Mutant 1,3-Propanediol Dehydrogenase
; FILE REFERENCE: GC580-2
; CURRENT APPLICATION NUMBER: US/09/570,778A
; CURRENT FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/134,868
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 387
; TYPE: PRF
; ORGANISM: Escherichia blattae
US-09-570-778A-8
```

```
Query Match          70.5%; Score 31; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 OYLRALGK 9
        1:11111:1
DB      342 QHLRELGVK 350
```

```
RESULT 14
US-09-570-778A-9
; Sequence 9, Application US/09570778A
; Patent No. 6468773
; GENERAL INFORMATION:
; APPLICANT: Trimbur, Donald E.
; APPLICANT: Whited, Gregory M.
; APPLICANT: Selfonova, Olga V.
; TITLE OF INVENTION: Mutant 1,3-Propanediol Dehydrogenase
; FILE REFERENCE: GC580-2
; CURRENT APPLICATION NUMBER: US/09/570,778A
; CURRENT FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/134,868
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 387
; TYPE: PRF
; ORGANISM: Escherichia blattae
US-09-570-778A-9
```

```
Query Match          70.5%; Score 31; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 OYLRALGK 9
        1:11111:1
DB      342 QHLRELGVK 350
```

```
RESULT 15
US-09-347-801-4
; Sequence 4, Application US/09347801
; Patent No. 6262345
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lee, Jian Ming
; TITLE OF INVENTION: Plant Protein Kinases
; FILE REFERENCE: BB-1171
; CURRENT APPLICATION NUMBER: US/09/347,801
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/092,438
; EARLIER FILING DATE: JULY 10, 1998
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Microsoft Office 97
```

```
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRF
; ORGANISM: Oryza sativa
US-09-347-801-4
```

```
Query Match          70.5%; Score 31; DB 4; Length 623;
Best Local Similarity 77.8%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 OYLRALGK 9
        11111111
DB      453 OYLRATPLK 461
```

Search completed: July 11, 2003, 12:02:08
Job time : 2.60915 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 2.63825 seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-5
Perfect score: 44
Sequence: 1 QYLRALGLK 9

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	US-10-091-442-5	Sequence 5, Appli
2	44	100.0	9	US-09-140-719-5	Sequence 5, Appli
3	44	100.0	380	US-10-091-442-34	Sequence 34, Appli
4	44	100.0	380	US-09-140-719-34	Sequence 34, Appli
5	34	77.3	350	US-09-847-208-53	Sequence 53, Appli
6	32	72.7	1465	US-10-083-357-1310	Sequence 1310, Ap
7	32	72.7	327	US-09-815-242-11463	Sequence 11463, A
8	32	72.7	327	US-09-815-242-11625	Sequence 11625, A
9	32	72.7	342	US-09-815-242-12104	Sequence 12104, A
10	31	70.5	292	US-10-023-282-1116	Sequence 1116, Ap
11	31	70.5	387	US-10-037-677-4	Sequence 4, Appli
12	31	70.5	387	US-09-991-138-2	Sequence 2, Appli
13	31	70.5	387	US-09-991-138-7	Sequence 7, Appli
14	31	70.5	387	US-09-991-138-8	Sequence 8, Appli
15	31	70.5	387	US-09-991-138-9	Sequence 9, Appli
16	31	70.5	424	US-10-179-784-3	Sequence 3, Appli
17	31	70.5	447	US-10-156-761-10914	Sequence 10914, A
18	31	70.5	478	US-10-128-714-3188	Sequence 3188, Ap
19	31	70.5	478	US-10-128-714-8188	Sequence 8188, Ap

20	31	70.5	623	US-09-854-731-4	Sequence 4, Appli
21	31	70.5	625	US-09-854-731-18	Sequence 18, Appli
22	31	70.5	905	US-10-156-761-14694	Sequence 14694, A
23	31	70.5	2548	US-09-851-682A-1	Sequence 1, Appli
24	30	68.2	328	US-09-895-913A-134	Sequence 134, App
25	30	68.2	331	US-10-147-874-2	Sequence 2, Appli
26	30	68.2	387	US-09-308-207-37	Sequence 37, Appli
27	30	68.2	387	US-09-991-138-10	Sequence 10, Appli
28	30	68.2	387	US-09-991-138-11	Sequence 11, Appli
29	30	68.2	449	US-09-815-242-11349	Sequence 11349, A
30	30	68.2	512	US-10-121-032-26	Sequence 26, Appli
31	30	68.2	512	US-10-093-037-26	Sequence 26, Appli
32	30	68.2	752	US-09-764-853-641	Sequence 641, App
33	30	68.2	1094	US-10-156-761-11435	Sequence 11435, A
34	30	68.2	1291	US-10-156-761-14161	Sequence 14161, A
35	30	68.2	1327	US-09-972-115A-8	Sequence 8, Appli
36	30	68.2	1327	US-09-841-835-2	Sequence 2, Appli
37	29	65.9	60	US-10-156-761-12478	Sequence 12478, A
38	29	65.9	61	US-09-738-626-4102	Sequence 4102, Ap
39	29	65.9	148	US-10-101-464A-553	Sequence 553, App
40	29	65.9	149	US-09-764-864-1171	Sequence 1171, Ap
41	29	65.9	214	US-10-101-464A-129	Sequence 129, App
42	29	65.9	214	US-09-864-636A-164	Sequence 164, App
43	29	65.9	223	US-09-892-398-22	Sequence 22, Appli
44	29	65.9	315	US-10-191-513A-16	Sequence 16, Appli
45	29	65.9	339	US-09-815-242-5446	Sequence 5446, Ap

ALIGNMENTS

RESULT 1
US-10-091-442-5
Sequence 5, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IWASA, Fuyuki
TSURUOKA, No. US20020164711A1
NAKAZATO, Hiroshi
MIURA, Kenji
ISHIDA, No. US20020164711A1
KURIHARA, Tatsuya
YAMACHI, KOZO
YAMAGUCHI, No. US20020164711A1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-091-442-5

Query Match 100.0%; Score 44; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
Db 1 QYLRALGLK 9

RESULT 2
US-09-140-719-5
Sequence 5, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
TITLE OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-5

Query Match 100.0%; Score 44; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
Db 1 QYLRALGLK 9

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
TITLE OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-MAR-2002
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 44; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUBOUOKA, No. US20010026931A1uo
APPLICANT: MAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 44; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 289 QYLRALGK 297

RESULT 5
US-09-847-208-53
Sequence 53, Application US/09847208
Publication No. US20030082190A1
GENERAL INFORMATION:
APPLICANT: Saxon, Andrew
APPLICANT: Zhang, Ke
APPLICANT: Zhu, Daocheng
TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
FILE REFERENCE: UC67.002A
CURRENT APPLICATION NUMBER: US/09/847,208
CURRENT FILING DATE: 2001-05-01
NUMBER OF SEQ ID NOS: 177
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 53
LENGTH: 350
TYPE: PRT
ORGANISM: Candida albicans (Yeast)
US-09-847-208-53

Query Match 77.3%; Score 34; DB 9; Length 350;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 191 QYLRALGK 199

RESULT 6
US-10-083-357-1310
Sequence 1310, Application US/10083357
Publication No. US20030054370A1
GENERAL INFORMATION:
APPLICANT: Qiantong Zeng et al.
TITLE OF INVENTION: Systemic Discovery of New Genes
FILE REFERENCE: 032796-090
CURRENT APPLICATION NUMBER: US/10/083,357
CURRENT FILING DATE: 2002-02-27
NUMBER OF SEQ ID NOS: 1346
SEQ ID NO 1310
LENGTH: 1465
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1310

Query Match 75.0%; Score 33; DB 9; Length 1465;
Best Local Similarity 66.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYLRALGK 9
DB 1094 QYLRALGK 1102

RESULT 7
US-09-815-242-11463
Sequence 11463, Application US/09815242
Patent No. US20020061569A1

```

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11463
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11463

Query Match      72.7%; Score 32; DB 10; Length 327;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 QYLRALGLK 9
DB      313 EYLNALGLQ 321

RESULT 8
US-09-815-242-11625
; Sequence 11625, Application US/09815242
; Patient No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12104
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12104

Query Match      72.7%; Score 32; DB 10; Length 342;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11625
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11625

Query Match      72.7%; Score 32; DB 10; Length 327;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 QYLRALGLK 9
DB      313 EYLNALGLQ 321

RESULT 9
US-09-815-242-12104
; Sequence 12104, Application US/09815242
; Patient No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12104
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12104

Query Match      72.7%; Score 32; DB 10; Length 342;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```
Publication No. US20030092893A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: p7007P1
CURRENT APPLICATION NUMBER: US/10/023,282
EARLIER FILING DATE: 2001-12-20
EARLIER APPLICATION NUMBER: 09/205,258
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1116
LENGTH: 292
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (11)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (15)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (35)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (36)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (39)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (40)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (45)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (91)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (255)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (256)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (257)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (258)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-1116

Query Match          70.5%; Score 31; DB 9; Length 292;
```

Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYLRALGK 7
Db 280 QHLRELGVK 286

RESULT 11
US-10-037-677-4
; Sequence 4, Application US/10037677
; Patent No. US20020173003A1
; GENERAL INFORMATION:
; APPLICANT: Schellenberger, Volker
; APPLICANT: Liu, Amy D.
; APPLICANT: Selifonova, Olga V.
; TITLE OF INVENTION: Directed Evolution of Microorganisms
; FILE REFERENCE: GC560
; CURRENT APPLICATION NUMBER: US/10/037,677
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/314,847
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 387
; TYPE: PRT
; ORGANISM: E. blatte
US-10-037-677-4

Query Match 70.5%; Score 31; DB 9; Length 387;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
Db 342 QHLRELGVK 350

RESULT 12
US-09-991-138-2
; Sequence 2, Application US/09991138
; Publication No. US20030040091A1
; GENERAL INFORMATION:
; APPLICANT: Trimbur, Donald E.
; APPLICANT: Whited, Gregory M.
; APPLICANT: Selifonova, Olga V.
; TITLE OF INVENTION: Mutant 1,3-Propanediol Dehydrogenase
; FILE REFERENCE: GC580-2D1
; CURRENT APPLICATION NUMBER: US/09/991,138
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/570,778
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/134,868
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Escherichia blattae
US-09-991-138-2

Query Match 70.5%; Score 31; DB 9; Length 387;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
Db 342 QHLRELGVK 350

RESULT 13

US-09-991-138-7
; Sequence 7, Application US/09991138
; Publication No. US20030040091A1
; GENERAL INFORMATION:
; APPLICANT: Trimbur, Donald E.
; APPLICANT: Whited, Gregory M.
; APPLICANT: Selifonova, Olga V.
; TITLE OF INVENTION: Mutant 1,3-Propanediol Dehydrogenase
; FILE REFERENCE: GC580-2D1
; CURRENT APPLICATION NUMBER: US/09/991,138
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/570,778
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/134,868
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Escherichia blattae
US-09-991-138-7

Query Match 70.5%; Score 31; DB 9; Length 387;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
Db 342 QHLRELGVK 350

RESULT 14
US-09-991-138-8
; Sequence 8, Application US/09991138
; Publication No. US20030040091A1
; GENERAL INFORMATION:
; APPLICANT: Trimbur, Donald E.
; APPLICANT: Whited, Gregory M.
; APPLICANT: Selifonova, Olga V.
; TITLE OF INVENTION: Mutant 1,3-Propanediol Dehydrogenase
; FILE REFERENCE: GC580-2D1
; CURRENT APPLICATION NUMBER: US/09/991,138
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/570,778
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/134,868
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Escherichia blattae
US-09-991-138-8

Query Match 70.5%; Score 31; DB 9; Length 387;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
Db 342 QHLRELGVK 350

RESULT 15
US-09-991-138-9
; Sequence 9, Application US/09991138
; Publication No. US20030040091A1
; GENERAL INFORMATION:
; APPLICANT: Trimbur, Donald E.
; APPLICANT: Whited, Gregory M.
; APPLICANT: Selifonova, Olga V.

; TITLE OF INVENTION: Mutant 1,3-Propanediol Dehydrogenase
; FILE REFERENCE: GC580-2D1
; CURRENT APPLICATION NUMBER: US/09/991,138
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 09/570,778
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/134,868
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Escherichia blattae
; US-09-991-138-9

Query Match 70.5%; Score 31; DB 9; Length 387;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Caps 0;

QY 1 QYLRALGLK 9
|:|:|:|:|
Db 342 QHLELGLVK 350

Search completed: July 11, 2003, 12:37:33
Job time : 2.63825 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 1.81497 Seconds

(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-5

Perfect score: 44

Sequence: 1 OYRALGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	79.5	164	2 I40526	hypothetical prote
2	35	79.5	270	2 B69781	multidrug-efflux t
3	34	77.3	80	2 F72303	glutaredoxin t
4	34	77.3	187	2 T06581	probable deoxyribo
5	34	77.3	350	1 S52153	alcohol dehydrogen
6	34	77.3	406	2 C75537	conserved hypotet
7	33	75.0	127	2 B64486	hypothetical prote
8	33	75.0	412	2 B81356	ankyrin repeat-con
9	33	75.0	413	2 E70661	probable PE protei
10	33	75.0	1139	2 T08421	resistance protein
11	33	75.0	1465	2 S13262	TYB protein - yeas
12	33	75.0	1467	2 PC1353	TYB protein - yeas
13	33	75.0	1802	2 S52611	TYB protein - yeas
14	33	75.0	1803	2 S56894	TYB protein - yeas
15	32	72.7	239	2 E84332	hypothetical prote
16	32	72.7	261	2 S20610	calpastatin - mous
17	32	72.7	274	2 S59378	hypothetical prote
18	32	72.7	309	2 C83886	hypothetical prote
19	32	72.7	337	2 H64716	ABC transporter, A
20	32	72.7	337	2 F71800	ABC transporter, A
21	32	72.7	342	2 A82968	alcohol dehydrogen
22	32	72.7	342	2 T34045	hypothetical prote
23	32	72.7	373	2 AD3470	alcohol dehydrogen
24	32	72.7	408	2 T33939	probable alcohol d
25	32	72.7	434	2 B64376	hypothetical prote
26	32	72.7	447	2 A72358	conserved hypotet
27	32	72.7	487	2 AD0818	probable exported
28	32	72.7	509	2 JC5651	N-acetylglucosamin
29	32	72.7	516	2 T24644	hypothetical prote

30	32	72.7	523	2 F82254	PTS system, N-acet
31	32	72.7	603	2 S15074	calpastatin - rat
32	32	72.7	625	2 G96976	probable periplasm
33	32	72.7	654	2 T10772	calpastatin - rat
34	32	72.7	1150	2 T40395	lim domain protein
35	32	72.7	1740	2 T43215	ribonucleotide red
36	31	70.5	107	2 F72059	hypothetical prote
37	31	70.5	107	2 E86565	hypothetical prote
38	31	70.5	127	2 T17431	hypothetical prote
39	31	70.5	240	2 T35972	hypothetical prote
40	31	70.5	342	2 A97201	hemol permease (lm
41	31	70.5	349	2 A97124	RECA recombinase,
42	31	70.5	352	2 A24648	alcohol dehydrogen
43	31	70.5	359	2 T36443	probable penicilli
44	31	70.5	375	2 AG2565	hypothetical prote
45	31	70.5	380	2 H82566	gunn protein Xf236

ALIGNMENTS

RESULT 1
I40526
hypothetical protein ypoC - Bacillus subtilis
N:Alternate names: hypothetical protein Y (ponA operon); jooC protein
C:Species: Bacillus subtilis
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text-change 20-Jun-2000
C:Accession: I40526; I40530; E69939
R:Brund, C.; Sorokin, A.; Serror, P.; Ehrlich, S.D.
Microbiology 141, 321-322, 1995
A:Title: Nucleotide sequence of the Bacillus subtilis dnaD gene.
A:Reference number: I40523; MUID:95219085; PMID:7704260
A:Accession: I40526
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <RES>
A:Cross-references: EMBL:U11289; NID:9533096; PIDN:AAA80006.1; PID:9533100
R:Popham, D.L.; Setlow, P.
J. Bacteriol. 177, 326-335, 1995
A:Title: Cloning, nucleotide sequence, and mutagenesis of the Bacillus subtilis ponA
A:Reference number: I40527; MUID:95113769; PMID:7814321
A:Accession: I40530
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <RES>
A:Cross-references: EMBL:U11883; NID:9520536; PIDN:AAA64948.1; PID:9520542
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.: Bron, S.; Brulliet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.: Ehrlich, S.D.; Emerson, P.T.; Ertlan, K.D.; Erlington, J.; Fabret, C.; Ferraril,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galliz, A.; Gal
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Huilo, M
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogihara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scari
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Seliguchi, J.; Sekowska, A.; Se
akench, M.; Tamakoshi, A.; Tanaka, T.; Terpsita, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipal, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69939
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-164 <KUN>
A:Cross-references: GB:Z99115; GB:AL009126; NID:92634478; PIDN:CAB14149.1; PID:926346
A:Experimental source: strain 168
C:Genetics:
A:Gene: ypoC; jooC
C:Superfamily: Bacillus subtilis hypothetical protein ypoC
Query Match 79.5% ; Score 35; DB 2; Length 164;

Best Local Similarity 77.8%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
| | | | |
Db 13 QYLRSLGOK 21

RESULT 2

B69781

multidrug-efflux transporter regulator homolog ydfl - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002

C:Accession: B69781

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

lechi, J.; Harwood, C.R.; Hentaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogn, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauder, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipit, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: B69781

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-270 <KUN>

A:Cross-references: GB:599106; GB:AL009126; NID:92632653; PIDN:CAB12353.1; PID:e1182512;

A:Experimental source: strain 168

C:Genetics:

A:Gene: ydfl

C:Superfamily: transcription activator bmrR

Query Match 79.5%; Score 35; DB 2; Length 270;

Best Local Similarity 87.5%; Pred. No. 15;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGL 8
| | | | |
Db 56 KYLRALGL 63

RESULT 3

F72303

glutaredoxin - Thermotoga maritima (strain MS8)

C:Species: Thermotoga maritima

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: F72303

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garratt, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72303

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-80 <ARN>

A:Cross-references: GB:AE001764; GB:AE000512; NID:94981561; PIDN:AMD36108.1; PID:9498157

A:Experimental source: strain MS8

C:Genetics:

A:Gene: TM1031

C:Superfamily: Clostridium pasteurianum probable glutaredoxin

C:Keywords: electron transfer

Query Match 77.3%; Score 34; DB 2; Length 80;

Best Local Similarity 66.7%; Pred. No. 7;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
| | | | |
Db 22 EYFRSLGOK 30

RESULT 4

T06581

probable deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - tomato (fragment)

N:Alternate names: cryptochrome 2; DNA photolyase; photoreactivating enzyme

C:Species: Lycopersicon esculentum (tomato)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000

C:Accession: T06581

R:Kolksaas, U.; Schneider-Poetsch, H.A.

submitted to the EMBL data library, August 1997

A:Description: The evolution of cryptochrome genes.

A:Reference number: Z15773

A:Accession: T06581

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-187 <KOL>

A:Cross-references: EMBL:AD000695; PIDN:CAA04247.1

C:Genetics:

A:Gene: CYP2

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: deoxyribodipyrimidine photo-lyase

C:Keywords: carbon-carbon lyase

Query Match 77.3%; Score 34; DB 2; Length 187;

Best Local Similarity 75.0%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YLRALGLK 9
| | | | |
Db 147 YLRALGLR 154

RESULT 5

S52153

alcohol dehydrogenase (EC 1.1.1.1) - yeast (Candida albicans)

N:Alternate names: 40K allergen

C:Species: Candida albicans

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Dec-2000

C:Accession: S63781; A61504; S52153

R:Berram, G.; Swoboda, R.K.; Gooday, G.W.; Gow, N.A.R.; Brown, A.J.P.

Yeast 12, 115-127, 1996

A:Title: Structure and regulation of the Candida albicans ADH1 gene encoding an immun

A:Reference number: S63781; MUID:96287648; PMID:8686375

A:Accession: S63781

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <BE2>

A:Cross-references: EMBL:X01694; NID:9608689; PIDN:CAA57342.1; PID:9608690

A:Note: only a part of the coding sequence is given

R:Shen, H.D.; Choo, K.B.; Lee, H.H.; Hsieh, J.C.; Lin, W.L.; Lee, W.R.; Han, S.H.

Clin. Exp. Allergy 21, 675-681, 1991

A:Title: The 40-kilodalton allergen of Candida albicans is an alcohol dehydrogenase:

A:Reference number: A61504; MUID:92136159; PMID:1777830

A:Accession: A61504

A:Molecule type: mRNA

A:Residues: 104-139, 'T', 'A', '228-312', 'S' <SHE>

A:Cross-references: GB:X81694; NID:9608689

A:Note: sequence extracted from NCBI backbone (NCBIN:80403, NCBI:P:80409)

C:Genetics:

A:Gene: ADH1

C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

C:Keywords: alcohol metabolism; NAD; oxidoreductase; zinc

F:31-338/Domain: long-chain alcohol dehydrogenase homology <LADH>

F:46,69,156/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted

Query Match 77.3%; Score 34; DB 1; Length 350;

Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 OYLRALGLK 9
Db 191 OYLRALGLR 199

RESULT 6

C75537

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: C75537

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1;

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: C75537

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-406 <WHI>

A:Cross-references: GB:AE001890; GB:AE000513; NID:96457959; PIDN:AF09871.1; PID:9645796

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0290

A:Map position: 1

C:Superfamily: tetracycline resistance protein

Query Match 77.3%; Score 34; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 36;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 YLRALGL 8
Db 33 YLRALGL 39

RESULT 7

B64486

hypothetical protein MJ1491 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: B64486

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: B64486

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-127 <BUU>

A:Cross-references: GB:U67590; GB:L77117; NID:q1592126; PIDN:AAB99510.1; PID:q1592128; T

C:Genetics:

A:Map position: FOR1465677-1466060

A:Start codon: GTC

C:Superfamily: Methanococcus jannaschii hypothetical protein MJ1491

Query Match 75.0%; Score 33; DB 2; Length 127;
Best Local Similarity 87.5%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 YLRALGLK 9
Db 83 YLRALGLK 90

RESULT 8

B81356
ankyrin repeat-containing probable periplasmic protein Cj0834c [imported] - Campyloba

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: B81356

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chli

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: B81356

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-412 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:96968128; PIDN:CAB73099.1; PID:9696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0834c

Query Match 75.0%; Score 33; DB 2; Length 412;
Best Local Similarity 75.0%; Pred. No. 58;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 YLRALGLK 9
Db 393 YLRALGLK 400

RESULT 9

E70661

probable PE protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: E70661

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

Comor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:96295987; PMID:9634230

A:Accession: E70661

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-413 <COU>

A:Cross-references: GB:Z83860; GB:AL123456; NID:q3261681; PIDN:CAB06157.1; PID:q17812

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: PE

C:Superfamily: Phaeosolus glycine-rich cell wall protein 1.8

Query Match 75.0%; Score 33; DB 2; Length 413;
Best Local Similarity 77.8%; Pred. No. 58;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 OYLRALGLK 9
Db 400 OYLRALGLK 408

RESULT 10

T08421

resistance protein homolog RGC2a - garden lettuce

N:Alternate names: resistance protein candidate

C:Species: Lactuca sativa (garden lettuce)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999

C:Accession: T08421

R:Shen, K.A.; Meyers, B.C.; Faridi, N.; Chin, D.B.; Stelly, D.M.; Michalmore, R.W.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z16412

A:Accession: T08421

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1139 <SHE>
 A:Cross-references: EMBL:AF017752; NID:g2852685; PID:g2852686
 C:Genetics:
 A:Gene: RGC2a
 A:Introns: 892/3

Query Match 75.0%; Score 33; DB 2; Length 1139;
 Best Local Similarity 55.6%; Pred. No. 1.6e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 |||:|:|
 Db 31 QYMRMGK 39

RESULT 11

S31262
 Tyb protein - yeast (*Saccharomyces cerevisiae*) retrotransposon Ty4 (fragment)
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 18-Jun-1993 #sequence_revision 25-Apr-1997 #text_change 26-May-2000
 C:Accession: S31262

R:Jancetky, B.; Lehle, L.
 J. Biol. Chem. 267, 19798-19805, 1992
 A:Title: Ty4, a new retrotransposon from *Saccharomyces cerevisiae*, flanked by tau-element
 A:Reference number: S31261; MUID:93015829; PMID:1328182
 A:Accession: S31262

A:Molecule type: DNA
 A:Residues: 1-1465 <JAN>
 A:Cross-references: EMBL:S46865
 C:Genetics:
 A:Mobile element: retrotransposon Ty4
 C:Superfamily: Tyb protein

Query Match 75.0%; Score 33; DB 2; Length 1465;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 |||:|:|
 Db 1094 QYLRGIGLK 1102

RESULT 12

PC1253
 Tyb protein - yeast (*Saccharomyces cerevisiae*) retrotransposon Ty4 (fragment)
 N:Contains: integrase; proteinase; reverse transcriptase; RNase H
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-May-2000
 C:Accession: PC1253; S27465

R:Stucka, R.; Schwarzlose, C.; Lochmüller, H.; Haecker, U.; Feldmann, H.
 Gene 122, 119-128, 1992
 A:Title: Molecular analysis of the yeast Ty4 element: Homology with Ty1, copia, and plant
 A:Reference number: JCI482; MUID:93083972; PMID:1333437
 A:Accession: PC1253

A:Molecule type: DNA
 A:Residues: 1-1467 <STU>
 R:Stucka, R.; Schwarzlose, C.; Lochmüller, H.; Hcker, U.; Feldmann, H.
 submitted to the EMBL Data Library, May 1992
 A:Description: Molecular analysis of the yeast Ty4 element: homology with Ty1, copia, and
 A:Reference number: S27465
 A:Accession: S27465

A:Molecule type: DNA
 A:Residues: 1-1467 <STW>
 A:Cross-references: EMBL:M94164
 C:Genetics:
 A:Gene: TY4B
 A:Mobile element: retrotransposon Ty4
 C:Superfamily: Tyb protein

Query Match 75.0%; Score 33; DB 2; Length 1467;
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 |||:|:|
 Db 1094 QYLRGIGLK 1102

RESULT 13

S52611
 Tyb protein - yeast (*Saccharomyces cerevisiae*) retrotransposon Ty4
 N:Alternate names: protein YHL008w-a
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 05-May-1995 #sequence_revision 19-Oct-1995 #text_change 23-Mar-2001
 C:Accession: S52611

R:Favell, T.
 submitted to the EMBL Data Library, June 1994
 A:Description: The sequence of S. *cerevisiae* cosmid L5018.
 A:Reference number: S46798
 A:Accession: S52611

A:Molecule type: DNA
 A:Residues: 1-1802 <FAV>
 A:Cross-references: EMBL:U11581
 C:Genetics:
 A:Map position: 8L
 A:Mobile element: retrotransposon Ty4
 C:Superfamily: Tyb protein

Query Match 75.0%; Score 33; DB 2; Length 1802;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 |||:|:|
 Db 1431 QYLRGIGLK 1439

RESULT 14

S56894
 Tyb protein - yeast (*Saccharomyces cerevisiae*) retrotransposon Ty4.JL
 N:Alternate names: protein J0780; protein YJL113w
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 08-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 17-Mar-2000
 C:Accession: S56894

R:Czepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
 submitted to the Protein Sequence Database, September 1995
 A:Reference number: S56894
 A:Accession: S56894
 A:Molecule type: DNA
 A:Residues: 1-1803 <CZJ>
 A:Cross-references: EMBL:Z49389
 C:Genetics:
 A:Map position: 10L
 A:Mobile element: retrotransposon Ty4.JL
 C:Superfamily: Tyb protein

Query Match 75.0%; Score 33; DB 2; Length 1803;
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 |||:|:|
 Db 1432 QYLRGIGLK 1440

RESULT 15

E84232
 hypothetical protein Vng0750c [imported] - *Halobacterium* sp. NRC-1
 C:Species: *Halobacterium* sp. NRC-1
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: E84232

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky
 Jung, K.H.; Alm, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;

A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: E84232
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <STO>
A:Cross-references: GB:AE004437; NID:g10580329; PIDN:AAG19225.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNG0750C

Query Match 72.7%; Score 32; DB 2; Length 239;
Best Local Similarly 85.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Caps 0;

OY 2 YLRALGL 8
| | | | | | | |
Db 6 YLRALGL 12

Search completed: July 11, 2003, 12:00:38
Job time : 3.91497 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 0.860707 Seconds

(Without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-5

Sequence: 1 QYLRALGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	380	1 SPB7_HUMAN	O75635 homo sapien
2	35	79.5	164	1 PPOC_BACSU	P39789 bacillus su
3	34	77.3	350	1 ADH1_CANAL	P43067 candida alb
4	33	75.0	127	1 YEG1_MERJA	O58886 methanococc
5	33	75.0	1803	1 YJL3_YEAST	P47024 saccharomyc
6	32	72.7	434	1 Y610_MERJA	O58027 methanococc
7	32	72.7	487	1 YFGC_SALTY	O6xg75 salmonella
8	32	72.7	654	1 ICAL_RAT	P23121 rattus norv
9	32	72.7	788	1 ICAL_MOUSE	P51125 mus musculu
10	31	70.5	349	1 RECA_CLOAB	P58254 clostridium
11	31	70.5	352	1 ADH3_EMENI	P07734 emeritella
12	31	70.5	355	1 ISPG_THETN	P07734 thermococ
13	31	70.5	383	1 HIS2_AQUAE	O67223 aquifex aeo
14	31	70.5	403	1 DACA_ECOLI	P04287 escherichia
15	31	70.5	619	1 GIDA_CAMEJ	O99na7 campylobact
16	31	70.5	670	1 DML1_BACST	O67703 bacillus st
17	31	70.5	960	1 CHP1_SCHRO	Q10103 schizosacch
18	30	68.2	262	1 V243_FOWPV	O91427 fowlpox vir
19	30	68.2	325	1 NARI_MOUSE	O60935 mus musculu
20	30	68.2	327	1 NARI_HUMAN	P52961 homo sapien
21	30	68.2	331	1 GLK_CAUCR	O9a6n3 cauldobacter
22	30	68.2	336	1 XYLZ_PSEPU	P23101 pseudomonas
23	30	68.2	345	1 H10W_BOVIN	P10950 bos taurus
24	30	68.2	375	1 ADH4_KLULA	P49385 kluyveromyc
25	30	68.2	387	1 DHAT_CITFR	P45513 citrobacter
26	30	68.2	387	1 ISCS_BUCAP	O59477 klebsiella
27	30	68.2	404	1 ISCS_BUCAP	O59477 klebsiella
28	30	68.2	419	1 LMB2_YERPE	O8rhp0 yersinia pe
29	30	68.2	449	1 MURC_HELPJ	O9a112 helicobacte
30	30	68.2	449	1 MURC_HELPJ	O9a112 helicobacte
31	30	68.2	470	1 ROS2_MOUSE	O62340 mus musculu
32	30	68.2	992	1 FRO_LEPAU	O01115 leptosphaer
33	30	68.2	1327	1 TWK1_HUMAN	O95271 homo sapien

34	30	68.2	1363	1 XDH_EMENI	Q12553 emeritella
35	30	68.2	1473	1 OVOS_CHICK	P20740 gallus gall
36	30	68.2	1545	1 MRP2_HUMAN	O92887 homo sapien
37	30	68.2	1666	1 MYM1_MOUSE	O62233 mus musculu
38	30	68.2	2051	1 FAS1_YEAST	P07149 s fatly aci
39	29	65.9	60	1 RL30_STRCO	P46789 streptomyc
40	29	65.9	74	1 VSH2_MICCU	P33104 micrococcus
41	29	65.9	141	1 VSH2_HABPA	P36434 haemophilus
42	29	65.9	162	1 PHZB_PSEFL	O51788 pseudomonas
43	29	65.9	170	1 YBET_ECOLI	O57822 methanococc
44	29	65.9	184	1 YBET_ECOLI	P77296 escherichia
45	29	65.9	190	1 Y516_AERPE	O9yer5 aeropyrum p

ALIGNMENTS

RESULT 1
SPB7_HUMAN STANDARD: PRT: 380 AA.
ID O75635: 15-JUN-2002 (Rel. 41, Created)
AC 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Megsin (TP55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:97326116; PubMed-9182567;
RA Tsurumoto M., Tsurumoto N., Ishida N., Kurihara T., Iwasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsunagi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mitsuuchi H., Yamaguchi N.,
RT "Purification, CDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity."
RT J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE:98376492; PubMed-9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.,
RT "A megakaryocyte-predominant gene, megasin, is a new serpin upregulated in
RT J. Clin. Invest. 102:828-836(1998).
RL
CC -!- FUNCTION: Might function as an inhibitor of lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in megakaryocytes.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL: D88575; BA3123.1; -
CC EMBL: AF027866; AAC64506.1; -
CC HSSP: P05619; IHDE.
CC Genew: HGNC:13902; SERPINB7.
CC MIM: 603357; -
CC InterPro: IPR000215; Serpin.
CC Pfam: PF00079; serpin; 1.
CC SMART: SM00093; SERPIN; 1.
CC PROSITE: PS00284; SERPIN; 1.

KW Serpin; Serine protease inhibitor.
 FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CF605 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QYRALGK 9
 DB 289 QYRALGK 297

RESULT 2
 YPOC_BACSU
 ID YPOC_BACSU STANDARD; PRT; 164 AA.
 AC P39789;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypoC (ORF).
 GN YPOC OR JOOC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95219085; PubMed=7704260;
 RT "Nucleotide sequence of the Bacillus subtilis dnaD gene."
 RL Microbiology 141:321-322(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95113769; PubMed=7814321;
 RT "Cloning, nucleotide sequence, and mutagenesis of the Bacillus subtilis ponA operon, which codes for penicillin-binding protein (PBP) 1 and a PBP-related factor."
 RL J. Bacteriol. 177:326-335(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RT "Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S., Borris R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Conerton I.F., Cummings S.D., Daniel R.A., Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holseppel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Bianchard M., Klein C., Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kunano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Potwilk S., Prescott A.M., Prescan E., Puje P., Purnelle B., Rapoport G., Ray M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter P., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni K., Toso V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Vari A., Wamutit R., Medler E., Medler K., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yamamoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,

RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."
 RL Nature 390:249-256(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U11289; AAA80006.1; -
 DR EMBL; U11883; AAA64948.1; -
 DR EMBL; L47709; AAB38458.1; -
 DR EMBL; Z99115; CAB14149.1; -
 DR Sublist; BG10955; YPOC.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 164 AA; 19346 MW; 5BB082046B7674CB CRC64;

Query Match 79.5%; Score 35; DB 1; Length 164;
 Best Local Similarity 77.8%; Pred. No. 3.5;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYRALGK 9
 DB 13 QYRALGK 21

RESULT 3
 ADH1_CANAL
 ID ADH1_CANAL STANDARD; PRT; 350 AA.
 AC P43067;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 35, Last annotation update)
 DE Alcohol dehydrogenase I (Ec 1.1.1.1) (40 kDa allergen) (Allergen Can a 1) (Can a 1).
 GN ADH1 OR CAD.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96287648; PubMed=8686375;
 RT "Structure and regulation of the Candida albicans ADH1 gene encoding an immunogenic alcohol dehydrogenase."
 RL Yeast 12:115-127(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B311A;
 RA Pendrak M.L., Klotz S.A., Smith R.L.;
 RL Submitted (Oct-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 104-313 FROM N.A.
 RX MEDLINE=92136159; PubMed=1777830;
 RA Shen H.D., Choo K.B., Lee H.H., Hsieh J.C., Lin W.L., Lee W.R., Han S.H.;
 RT "The 40-kilodalton allergen of Candida albicans is an alcohol dehydrogenase: molecular cloning and immunological analysis using monoclonal antibodies."
 RL Clin. Exp. Allergy 21:675-681(1991).
 CC -1- CATALYTIC ACTIVITY: An alcohol + NAD(+) - an aldehyde or ketone + NADH.
 CC -1- COFACTOR: ZINC.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X81694; CAA57342.1; -
 DR EMBL: U15924; AAB53300.1; -
 DR COMPLEYEAST-2DPAGE; P43067; -
 DR InterPro: IPR002328; Adh_zinc.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; Adh_zinc; 1.
 KM Oxidoreductase; Zinc; NAD; Allergen.
 FT METAL 46 46 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 103 103 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 106 106 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 114 114 ZINC (SECOND ATOM) (BY SIMILARITY).
 FT METAL 156 156 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CONFLICT 39 39 H -> N (IN REF. 2).
 FT CONFLICT 53 53 R -> W (IN REF. 3).
 FT CONFLICT 140 140 A -> T (IN REF. 3).
 FT CONFLICT 212 212 F -> L (IN REF. 2).
 FT CONFLICT 227 227 D -> A (IN REF. 3).
 FT CONFLICT 313 313 R -> S (IN REF. 3).
 FT CONFLICT 327 327 D -> E (IN REF. 2).
 SQ SEQUENCE 350 AA; 36879 MW; 330FE2E70F42D634 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 350;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYRALGLK 9
 DB 191 QYRAMGLR 199

RESULT 4
 ID YE91_METJA STANDARD: PRT; 127 AA.
 AC Q58886;
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Hypothetical protein MJ1491.
 GN MJ1491.
 OS Methanococcus jannaschii.
 OC Archaea: Euryarchaeota: Methanococci: Methanococcales;
 CC Methanocaldococcaceae: Methanocaldococcus.
 OX NCBI_TaxID:2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE-96337999; PubMed-8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerelevage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
 RA Overbeek R., Kirkness E.F., Weisscock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Meldrum J.F., Fuhmann J.L., Nguyen D.,
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential)
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U67590; AAB99510.1; -
 DR TIGR: MJ1491; -
 KM Hypothetical protein: Transmembrane; Complete proteome.
 FT TRANSMEM 16 36 POTENTIAL.
 FT TRANSMEM 59 79 POTENTIAL.
 FT TRANSMEM 100 120 POTENTIAL.
 SQ SEQUENCE 127 AA; 13006 MW; 6D3FB594A7961C96 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 127;
 Best Local Similarity 87.5%; Pred. No. 7.1;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
 DB 83 YLSALGLK 90

RESULT 5
 ID YJL3_YEAST STANDARD: PRT; 1803 AA.
 AC P47024; P87192;
 DT 01-FEB-1996 (Rel. 33; Created)
 DT 30-MAY-2000 (Rel. 39; Last sequence update)
 DT 30-MAY-2000 (Rel. 39; Last annotation update)
 DE Transposon Ty4 207.7 kDa hypothetical protein.
 GN TY4B OR YJL113W OR J0780.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID:4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / FY1679;
 RX MEDLINE-97103775; PubMed-8948101;
 RA Cziepluch C., Kordes E., Pujol A., Jauniaux J.-C.;
 RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
 RT reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
 RT SPI10, GC014, Rpt1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,
 RT three remnant delta elements and a Ty4 transposon.";
 RL Yeast 12:1471-1474(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: 249389; CAA89409.1; -
 DR SGD: S0003649; YJL113W.
 DR InterPro: IPR001584; Rve.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF00665; rve; 1.
 DR SMART: SM00343; ZNF_C2HC; 1.
 KW Transposable element; Hypothetical protein.
 SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284ABD52D3 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 1803;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 QYRALGLK 9
 DB 1432 QYLANGGLK 1440

RESULT 6

Y610_METIA
ID 1610_METIA STANDARD; PRT: 434 AA.
AC Q58027;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0610.
GN M0610.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96371999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Utermack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Kloten H.-P., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
RA Klein M.D., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.;
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO C.ELEGANS D1054.13.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67509; AAB98603.1; -
DR TIGR; M00610; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 434 AA; 48683 MW; 09AFBC5819613EF5 CRC64;
QY Query Match 72.7%; Score 32; DB 1; Length 434;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 2 YLRALGLK 9
121 FLKALGLK 128
RESULT 7
YFGC_SALTY
ID YFGC_SALTY STANDARD; PRT: 487 AA.
AC Q8XG75;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yfgc precursor.
GN YFGC OR STM2494 OR STY2735.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium; STRAIN=LT2 / SGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McLelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Ngoylik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;
RT *Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhi; STRAIN=CR18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gara P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT *Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CR18.";
RL Nature 413:848-852(2001).
CC -1- SIMILARITY: CONTAINS 2 TPR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008812; AAL21388.1; -
DR DR Styene; AL627275; CAD02696.1; -
DR Styene; SG77777; YFGC.
DR InterPro; IPR001440; TPR.
KW Hypothetical protein; Signal; Repeat; TPR repeat; Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 487 POTENTIAL.
FT REPEAT 309 342 HYPOTHETICAL PROTEIN YFGC.
FT REPEAT 427 460 TPR 1.
FT REPEAT 427 460 TPR 2.
SQ SEQUENCE 487 AA; 53740 MW; 593FCFC8DBCCCE CRC64;
QY Query Match 72.7%; Score 32; DB 1; Length 487;
Best Local Similarity 55.6%; Pred. No. 44;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 1 QYRALGLK 9
74 QYINALGMR 82
RESULT 8
ICAL_RAT
ID ICAL_RAT STANDARD; PRT: 654 AA.
AC P27321;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calpain inhibitor (Calpastatin).
GN CAST.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91198152; PubMed=2015306;
RA Ishida S., Emori Y., Suzuki K.,
RT "Rat calpastatin has diverged primary sequence from other mammalian
RL calpastatins but retains functionally important sequences.";
RN [2]
RP SEQUENCE OF 1-129 FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=Fischer;

RX MEDLINE-92110395; PubMed-1730065;
 RA Lee W.J., Hatanaka M., Maki M.;
 RT "Multiple forms of rat calpastatin cDNA in the coding region of
 functionally unknown amino-terminal domain.";
 RL Blochum. Biophys. Acta 1129:251-253(1992).
 CC -1- FUNCTION: Specific inhibition of calpain (calcium-dependent
 cysteine protease). Plays a key role in postmortem tenderization
 of meat and have been proposed to be involved in muscle protein
 degradation in living tissue.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: A (SHOWN HERE), B AND C; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DOMAIN: HAS FOUR INHIBITORY DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE CALPASTATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X56729; CAA40053.1; -;
 DR EMBL: X62520; CAA44386.1; -;
 DR PIR: S15074; S15074.
 DR InterPro: IPR001259; Calpain_inhnb.
 DR Pfam: PF00748; Calpain_inhnb: 4.
 KM Repeat: Thiol protease inhibitor: Alternative splicing.
 FT REPEAT 149 201 INHIBITORY DOMAIN 1.
 FT REPEAT 282 334 INHIBITORY DOMAIN 2.
 FT REPEAT 392 445 INHIBITORY DOMAIN 3.
 FT REPEAT 529 582 INHIBITORY DOMAIN 4.
 FT VARSPPLIC 8 20 MISSING (IN ISOFORM C).
 FT VARSPPLIC 39 76 MISSING (IN ISOFORM B AND ISOFORM C).
 FT CONFLICT 102 102 K -> E (IN REF. 1).
 SQ SEQUENCE 654 AA: 71367 MW: 497089A198D819A CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 654;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 YLRALGLK 9
 Db 169 YLEALGIK 176
 RESULT 9
 ID ICAL_MOUSE STANDARD; PRT; 788 AA.
 AC P51125; O9RON1; O9OX03; O9OX04;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUN-2003 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calpain inhibitor (Calpastatin).
 DE CAST.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS MCS-A; MCS-B AND MCS-C).
 RC STRAIN=C3H; TISSUE=Muscle;
 RX MEDLINE-99333681; PubMed-10403772;
 RA Takano J., Kawamura T., Murase M., Hltoml K., Maki M.;
 RT "Structure of mouse calpastatin isoforms: implications of species-
 common and species-specific alternative splicing."
 RL Biochem. Biophys. Res. Commun. 260:339-345(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS TCAST1 AND TCAST2).
 RC STRAIN=CD-1;
 RA Li S., Goldberg E.;
 RT "Characterization of a membrane associated testis-specific
 calpastatin.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 84-363 FROM N.A. (ISOFORM MCS-B).
 RC STRAIN=BA1B/C; TISSUE=Heart;
 RX MEDLINE-92110395; PubMed-1730065;
 RA Lee W.J., Hatanaka M., Maki M.;
 RT "Multiple forms of rat calpastatin cDNA in the coding region of
 functionally unknown amino-terminal domain.";
 RL Blochum. Biophys. Acta 1129:251-253(1992).
 CC -1- FUNCTION: Specific inhibition of calpain (calcium-dependent
 cysteine protease). Plays a key role in postmortem tenderization
 of meat and have been proposed to be involved in muscle protein
 degradation in living tissue.
 CC -1- DOMAIN: HAS FOUR INHIBITORY DOMAINS.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: MCS-A (SHOWN HERE), MCS-B, MCS-
 C, TCAST1 AND TCAST2; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ISOFORM MCS-B IS THE MAJOR FORM IN ALL TISSUES
 EXAMINED. ISOFORM MCS-A ACCOUNTS FOR 5-10% IN TISSUES SUCH AS
 SKELETAL MUSCLE, LIVER AND BRAIN, AND 30% IN MYOBLASTS. ISOFORMS
 TCAST1 AND TCAST2 ARE TESTIS-SPECIFIC.
 CC -1- SIMILARITY: BELONGS TO THE CALPASTATIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AB026997; BAA84768.1; -;
 DR EMBL: AF190152; AAF25194.1; -;
 DR EMBL: AF190151; AAF25193.1; -;
 DR EMBL: X62519; CAA44385.1; -;
 DR MGD: MGI:1098236; Cast.
 DR InterPro: IPR001259; Calpain_inhnb.
 DR Pfam: PF00748; Calpain_inhnb: 2.
 KM Repeat: Thiol protease inhibitor: Alternative splicing.
 FT REPEAT 251 303 INHIBITORY DOMAIN 1.
 FT REPEAT 384 436 INHIBITORY DOMAIN 2.
 FT REPEAT 524 577 INHIBITORY DOMAIN 3.
 FT REPEAT 661 714 INHIBITORY DOMAIN 4.
 FT VARSPPLIC 1 341 MISSING (IN ISOFORM TCAST1).
 FT VARSPPLIC 1 357 MISSING (IN ISOFORM TCAST2).
 FT VARSPPLIC 92 110 MISSING (IN ISOFORM MCS-B AND ISOFORM
 MCS-C).
 FT VARSPPLIC 305 333 MISSING (IN ISOFORM MCS-C).
 FT VARSPPLIC 342 396 PDPSHVTSQAEQVK -> MGQFLSTFWGSPAAVWQEKLR
 EGERKAGETIPILDDHYICSEERHSGKH (IN
 ISOFORM TCAST1).
 FT KRVKEEVINDALALSLCTRPDPSPSHVTSQAEQVE
 -> MGQFLSTFWGSPAAVWQEKLRGKAGETIPILQ
 DH (IN ISOFORM TCAST2).
 SQ SEQUENCE 788 AA: 84921 MW: 28E4D3B4A68BFE9 CRC64;
 Query Match 72.7%; Score 32; DB 1; Length 788;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 YLRALGLK 9
 Db 271 YLEALGIK 278
 RESULT 10
 ID RECA_C10AB STANDARD; PRT; 349 AA.
 AC P58254;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE RecA protein (Recombinase A).

GN RECA OR CAC1815.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RM SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE-21359325; PubMed-11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
Gibson R.L., Lee H.M., Dubois J., Qiu D., Hill J., Wolf Y.I.,
Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
Bennett G.N., Koonin E.V., Smith D.R.;
RT Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 185:4823-4838(2001).
CC -1- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF
CC SINGLE-STRANDED DNA. THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED
CC DNA BY DUPLEX DNA, AND THE ATP-DEPENDENT HYBRIDIZATION OF
CC HOMOLOGOUS SINGLE-STRANDED DNAs. IT INTERACTS WITH LEXA CAUSING
CC ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RECA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AE007650; AAK79780.1; -
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR001553; RECA.
DR Pfam: PF00154; reca; 1.
DR PRINTS: PR00142; RECA.
DR PRODOM: PD000229; RecA; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00321; RECA_1; 1.
DR PROSITE: PS50162; RECA_2; 1.
DR PROSITE: PS50163; RECA_3; 1.
KM DNA damage: DNA recombination; SOS response; ATP-binding; DNA-binding;
KM Complete proteome.
FT NP_BIND 65 72 ATP (By similarity).
SQ SEQUENCE 349 AA; 38089 MW; ICB821EDFB3C855D CRC64;
Query Match 70.5%; Score 31; DB 1; Length 349;
Best Local Similarity 75.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 QYRALGL 8
DB 101 QYRALGLV 108
RESULT 11
ADH3_EMENT
ID ADH3_EMENT STANDARD; PRT; 352 AA.
AC P07754;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase III (EC 1.1.1.1) (ADH III).
GN ALDC OR ADH3.
OS *Escherichia nidulans* (Aspergillus nidulans).
OC Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; *Escherichia*.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86055769; PubMed=2998782;

RA McKnight L., Kato H., Upshall A., Parker M.D., O'Hara P.J., O'Hara S.;
RT "Identification and molecular analysis of a third *Aspergillus*
RT *nidulans* alcohol dehydrogenase gene.";
RL EMO J. 4:2093-2099(1985).
CC -1- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
CC NADH.
CC -1- COFACTOR: ZINC.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, X02764; CAA26541.1; -
DR PIR, A24648; A24648.
DR InterPro: IPR002328; Adh_zinc.
DR InterPro: IPR002085; Adh_zn_family.
DR Pfam, PF00107; adh_zinc; 1.
DR PROSITE, PS00059; ADH_ZINC; 1.
KM Oxidoreductase; zinc; NAD; Multigene family.
FT METAL 44 44 ZINC (Catalytic).
FT METAL 67 67 ZINC (Catalytic).
FT METAL 100 100 ZINC (SECOND ATOM).
FT METAL 103 103 ZINC (SECOND ATOM).
FT METAL 106 106 ZINC (SECOND ATOM).
FT METAL 114 114 ZINC (SECOND ATOM).
FT METAL 156 156 ZINC (Catalytic).
SQ SEQUENCE 352 AA; 37126 MW; AFDE78FE442E3144 CRC64;
Query Match 70.5%; Score 31; DB 1; Length 352;
Best Local Similarity 55.6%; Pred. No. 51;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 QYRALGL 9
DB 191 QYRALGLR 199
RESULT 12
ISPG_THETN
ID ISPG_THETN STANDARD; PRT; 355 AA.
AC ORR430;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase.
GN ISPG OR GCPE OR TREL400.
OS *Thermoanaerobacter tengcongensis*.
OC Bacteria; Firmicutes; Clostridia; *Thermoanaerobacteriales*;
OC *Thermoanaerobacteriaceae*; *Thermoanaerobacter*.
OX NCBI_TaxID=119072;
RN [1]
RM SEQUENCE FROM N.A.
RC STRAIN-M84 / JCM 11007;
RX MEDLINE-21992816; PubMed-11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of *T. tengcongensis* genome.";
RL Genome Res. 12:689-700(2002).
CC -1- FUNCTION: Converts 2C-methyl-D-erythritol 2,4-cyclodiphosphate
CC (ME-2,4cPP) into 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
CC (By similarity).
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; sixth step.
CC -1- SIMILARITY: BELONGS TO THE ISPG FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE013100; AAM24622.1; -
 KW Isoprene biosynthesis: Complete proteome.
 SQ SEQUENCE 355 AA; 39005 MW; ADDEF97DE5622A0 CRC64;

Query Match 70.5%; Score 31; DB 1; Length 355;
 Best Local Similarity 87.5%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 QYLRALGL 8
 Db 249 QILRALGL 256

RESULT 13
 HIS2_AOUAE STANDARD; PRT; 383 AA.
 AC O67223;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP phosphotransferase regulatory subunit.
 GN HIS2 OR HIS2 OR AQ_1155.
 OS Aquifex aeolicus.
 OC Bacteria: Aquificae: Aquificae (class): Aquificaceae; Aquificaceae;
 OC Aquifex.

NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.

RC MEDLINE=98196666; PubMed=9537320;
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus";
 RL Nature 392:353-358(1998).

CC -1- FUNCTION: May allow the regulation of ATP
 CC phosphotransferase activity by histidine (By similarity).
 CC -1- PATHWAY: Histidine biosynthesis; first step. Very important in the
 CC regulation of histidine metabolism.
 CC -1- SUBUNIT: Binds to hisG (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- MISCELLANEOUS: This function is generally fulfilled by the C-
 CC terminal part of hisG, which is missing in some bacteria such as
 CC this one.

CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC HIS2 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AE000726; AAC07188.1; -
 DR HSSP: P04804; IKM.

DR InterPro: IPR002106; AATRNA_1igaaseII.
 DR InterPro: IPR004517; His2_second.

DR InterPro: IPR002314; tRNA-synt.2b.
 DR Pfam: PF00587; tRNA-synt.2b; 1.

DR TIGRFAMs: TIGR00443; his2_second; 1.
 DR Hisididine biosynthesis: Complete proteome.

KW SEQUENCE 383 AA; 44349 MW; E3B449B5BF86C8D CRC64;

Query Match 70.5%; Score 31; DB 1; Length 383;
 Best Local Similarity 55.6%; Pred. No. 56;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QYLRALGLK 9
 Db 140 EFLKSLGLK 148

RESULT 14

DNA_ECOLI STANDARD; PRT; 403 AA.
 AC P04287; P77106;

DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Penicillin-binding protein 5 precursor (D-alanyl-D-alanine
 DE carboxypeptidase fraction A) (Ec 3.4.16.4) (DD-peptidase)
 DE (DD-carboxypeptidase) (PBP-5).
 GN DACC OR PIV OR B0632 OR 20777 OR ECS0670.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

NCBI_TaxID=562, 83334;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RC MEDLINE=88157719; PubMed=3279397;
 RA Broome-Smith J.K., Ioannidis I., Edelman A., Spratt B.G.;
 RT "Nucleotide sequences of the penicillin-binding protein 5 and 6 genes
 RT of Escherichia coli";
 RL Nucleic Acids Res. 16:1617-1617(1988).

RP SEQUENCE FROM N.A.
 RA Broome-Smith J.K.;
 RL Submitted (APR-1984) to the EMBL/Genbank/DBJ databases.

RP [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RC MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).

RP [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RC MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino Y., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yanai M., Horiiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map";
 RL DNA Res. 3:137-155(1996).

RP [5]
 RP SEQUENCE FROM N.A.
 RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
 RA Federspiel N., Hyman R., Kaiman S., Komp C., Kurd O., Lew H.,
 RA Lin D., Mamath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.

RP [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RC MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7,"
 RL Nature 409:529-533(2001).
 [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 0509952;
 RX MEDLINE-21156231; Pubmed-11238796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12,"
 RL DNA Res. 8:11-22(2001).
 [8]
 RP SEQUENCE OF 108-336 FROM N.A., AND MUTANT DACA11191.
 RX MEDLINE-84108878; Pubmed-6319180;
 RA Broome-Smith J.K., Spratt B.G.,
 RT "An amino acid substitution that blocks the deacylation step in the
 RT enzyme mechanism of penicillin-binding protein 5 of *Escherichia*
 RT coli,"
 RL FEBS Lett. 165:185-189(1984).
 [9]
 RP SEQUENCE OF 1-39 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-86058785; Pubmed-2316191;
 RA Takase I., Ishino F., Machi M., Kamata H., Doi M., Asoh S.,
 RA Matsuzawa H., Ohta T., Matsushashi M.,
 RT "Genes encoding two lipoproteins in the *leuS*-*dacA* region of the
 RT *Escherichia coli* chromosome,"
 RL J. Bacteriol. 169:5692-5699(1987).
 [10]
 RP SEQUENCE OF 30-57.
 RX MEDLINE-82187192; Pubmed-7042389;
 RA Maxam D.J., Amanuma H., Strominger J.L.,
 RT "Amino acid sequence homologues between *Escherichia coli* penicillin-
 RT binding protein 5 and class A beta-lactamases,"
 RL FEBS Lett. 139:159-163(1982).
 [11]
 RP MUTAGENESIS OF LYS-213.
 RX MEDLINE-92283850; Pubmed-1597468;
 RA Malhotra K.T., Nicholas R.A.,
 RT "Substitution of lysine 213 with arginine in penicillin-binding
 RT protein 5 of *Escherichia coli* abolishes D-alanine carboxypeptidase
 RT activity without affecting penicillin binding,"
 RL J. Biol. Chem. 267:11386-11391(1992).
 [12]
 RP DOMAINS.
 RX MEDLINE-93143736; Pubmed-8424800;
 RA van der Linden M.P.G., de Haan L., Keck W.,
 RT "Domain organization of penicillin-binding protein 5 from *Escherichia*
 RT coli analysed by C-terminal truncation,"
 RL Biochem. J. 289:593-598(1993).
 [13]
 RP FUNCTION: REMOVES C-TERMINAL D-ALANYL RESIDUES FROM SUGAR-PEPTIDE
 CC CELL WALL PRECURSORS.
 CC -1- CATALYTIC ACTIVITY: D-alanyl-D-alanine + H(2)O = 2 D-alanine.
 CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED. N-TERMINAL
 CC LIES IN THE PERIPLASMIC SPACE.
 CC -1- MISCELLANEOUS: THE MUTANT DACA1191 IS STILL BINDING PENICILLIN,
 CC BUT IS BLOCKED IN THE RELEASE OF THE BOUND PENICILLOYL MOIETY; THE
 CC MUTANT ALSO FAILS TO CATALYSE THE D-ALANINE CARBOXYPEPTIDASE
 CC REACTION AS THE HYDROLYSIS OF THE ACYL-ENZYME FORMED WITH
 CC SUBSTRATE IS ALSO BLOCKED AND THE ACYL-ENZYME ACCUMULATES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S11, ALSO KNOWN AS THE
 CC D-ALANYL-D-ALANINE CARBOXYPEPTIDASE 1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; X06479; CAA29774.1; -;
 CC DR EMBL; M18276; AAA24553.1; -;
 CC DR EMBL; AE000168; AAC73733.1; -;
 CC DR EMBL; D90703; BAA35275.1; -;
 CC DR EMBL; D90704; BAA35279.1; -;
 CC DR EMBL; U82598; AAB40832.1; ALT_INIT.
 CC DR EMBL; AE005242; AAC54966.1; -;
 CC DR EMBL; AP002552; BAB34093.1; -;
 CC DR EMBL; L07636; AAA66340.1; -;
 CC DR PIR; A03420; ZPECP5.
 CC DR PIR; A28536; A28536.
 CC DR PIR; D28387; D28387.
 CC DR MEROPS; S11.003; -;
 CC DR EC02DBASE; H040.5; 6TH EDITION.
 CC DR Ecogene; EG10201; dacA.
 CC DR InterPro; IPR001967; Ala/AlaCbpIase1.
 CC DR Pfam; PF00768; Peptidase_S11; 1.
 CC DR PRINTS; PR00725; DADACBPASE1.
 CC KW Hydrolyase; Carboxypeptidase; Peptidoglycan synthetase; Inner membrane;
 CC Cell wall; Signal; Complete proteome.
 CC FT SIGNAL 1 29
 CC FT CHAIN 30 403
 CC FT ACT_SITE 73 73
 CC FT VARIANT 134 134
 CC FT CONFLICT 94 94
 CC FT SEQUENCE 403 AA; 44444 MW; 7FAAB8E98452EF22 CRC64;
 CC
 CC Query Match 70.5%; Score 31; DB 1; Length 403;
 CC Best Local Similarity 75.0%; Pred. No. 59;
 CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 YLRALGK 9
 CC Db 165 YVNALGK 172
 CC
 CC RESULT 15
 CC GIDA_CAMME
 CC ID GIDA_CAMME STANDARD; PRT; 619 AA.
 CC AC Q9PNA7;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Glucose inhibited division protein A.
 CC GN GIDA OR CJ1188C.
 CC OS Campylobacter jejuni.
 CC OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 CC OC Campylobacter.
 CC OX NCBI_TaxID=197;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-NCTC 11168;
 CC RX MEDLINE-20150912; Pubmed-10688204;
 CC RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,
 CC Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 CC Jagers K., Karlyshev A.V., Moule S., Fallen M.J., Penn C.W.,
 CC Ouall M.A., Rajandream M.G., Rutherford K.M., van Vliet A.H.M.,
 CC Whitehead S., Barrett B.G.,
 RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*
 RT reveals hypervariable sequences,"
 RL Nature 403:665-668(2000).
 CC -1- FUNCTION: NOT KNOWN.
 CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: AL139077; CAB73442.1; -
 DR InterPro: IPR001327; FAD_pyr_redox.
 DR InterPro: IPR002218; GIDA.
 DR InterPro: IPR004416; GIDA_sub.
 DR InterPro: IPR002645; STAS.
 DR Pfam: PF01134; GIDA.1.
 DR PRINTS: PR00368; FADPNR.
 DR PRODOM: PD003738; GIDA.1.
 DR TIGRFAMS: TIGR00136; gida.1.
 DR PROSITE: PS01280; GIDA_1; 1.
 DR PROSITE: PS01281; GIDA_2; FALSE_NEG.
 KW Complete proteome.
 SO SEQUENCE 619 AA; 69172 MW; 18261C2019852AE2 CRC64;

Query Match 70.5%; Score 31; DB 1; Length 619;
 Best Local Similarity 75.0%; Pred. No. 90;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
 11: 1111
 DB 181 YLRALGLK 188

Search completed: July 11, 2003, 11:55:16
 Job time: 1.86071 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 3.92931 Seconds
(without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-5
Perfect score: 44
Sequence: 1 OYLALGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	79.5	270	16	P96690 Bacillus su
2	35	79.5	802	11	O8VE90 mus musculu
3	35	79.5	808	4	O96A78 homo sapien
4	34	77.3	80	16	O9X0C1 thermotoga
5	34	77.3	187	10	O2A020 lycopersico
6	34	77.3	361	3	O9HDD6 cryptococcu
7	34	77.3	363	2	O9AGZ2 ethiopia c
8	34	77.3	406	16	O9RXM0 delinococcu
9	34	77.3	635	10	O9LDJ5 lycopersico
10	34	77.3	768	12	O9IBW6 turkey heip
11	34	77.3	769	12	O9E6B8 turkey heip
12	34	77.3	1013	10	O9FXM9 arabidopsis
13	34	75.0	228	16	O8XK60 clostridium
14	33	75.0	301	16	O8R8F7 thermotoga
15	33	75.0	314	17	O8ZT63 pyrobaculum
16	33	75.0	348	5	O9GYX0 strongyloce

17	33	75.0	412	16	O9PE89 campylobact
18	33	75.0	413	16	P95236 mycobacteri
19	33	75.0	415	17	O96Z30 sulfolobus
20	33	75.0	1139	10	O48894 lactuca sat
21	33	75.0	1885	17	O8TDD1 bacterioph
22	33	72.7	80	9	O03961 methanosarc
23	32	72.7	154	5	O9GRR9 leishmania
24	32	72.7	176	11	O51555 ratius norv
25	32	72.7	194	5	O9V827 drosophila
26	32	72.7	201	12	O9DSU7 ascovirus d
27	32	72.7	239	17	O9HRD3 halobacteri
28	32	72.7	274	3	O06688 saccharomyc
29	32	72.7	295	17	O8T509 methanosarc
30	32	72.7	303	2	O93DS5 streptococc
31	32	72.7	309	16	O9KRM4 bacillus ha
32	32	72.7	311	17	O976M4 sulfolobus
33	32	72.7	327	2	O30552 helicobacte
34	32	72.7	327	16	O26096 helicobacte
35	32	72.7	327	16	O9ZJ34 helicobacte
36	32	72.7	342	5	O02093 caenorhabdi
37	32	72.7	342	16	O9HMD9 pseudomonas
38	32	72.7	373	16	O8YET1 brucella me
39	32	72.7	408	5	O9UAT1 caenorhabdi
40	32	72.7	422	16	O8Y266 raistonia s
41	32	72.7	447	16	O9WZ69 thermotoga
42	32	72.7	481	4	O9NUP7 homo sapien
43	32	72.7	506	16	O8RET1 fusobacteri
44	32	72.7	509	2	O32444 vibrio chol
45	32	72.7	516	5	O22287 caenorhabdi

ALIGNMENTS

RESULT 1
ID P96690 PRELIMINARY: PRT: 270 AA.
AC P96690:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE YDFL protein.
GN YDFL.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,
RA Broiliet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codan J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Dauter F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Euteneier K.D., Erlingsson J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Gallier N.,
RA Gilm S.V., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kletter-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

```

RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara B., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portelle D., Potwollit S., Prescott A.M.,
RA Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Roche E., Roche B., Rose M., Sadleir Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldó B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takanishi M., Tamakoshi A., Tanaka T., Terpsila P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A.,
RA Viari A., Wambuit R., Wedler E., Wedler H., Wietzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: AB001488; BAI19380.1; -
DR EMBL: Z99106; CAB12353.1; -
DR InterPro: IPR000551; HTH_Merr.
DR Pfam: PF00376; merr; 1.
DR SMART: SM00422; HTH_MERR; 1.
DR DNA-binding: Transcription regulation: Complete proteome.
KM KW
SQ SEQUENCE 270 AA; 31783 MW; 1C487674F59A6797 CRC64;

Query Match 79.5%; Score 35; DB 16; Length 270;
Best Local Similarity 87.5%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QYRAAGL 8
Db 56 KYRAAGL 63

RESULT 2
OY 08VE90 PRELIMINARY; PRT; 802 AA.
AC 08VE90;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 88.4 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC019504; AAH19504.1; -
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 12.
DR PRINTS: PR00320; GPROTEINBRPT.
DR SMART: SM00320; WD40; 13.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE: PS50082; WD_REPEATS_2; 9.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM KW
SQ SEQUENCE 802 AA; 88379 MW; EFC3CFB04245431B CRC64;

Query Match 79.5%; Score 35; DB 11; Length 802;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QYRAAGL 8

```

```

Db 666 KYRAAGL 673

RESULT 3
OY 096A78 PRELIMINARY; PRT; 808 AA.
AC 096A78;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 89.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LONG;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-CERVIX;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 13 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: BC014007; AAH14007.1; -
DR EMBL: BC010231; AAH10231.1; -
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 13.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE: PS50082; WD_REPEATS_2; 8.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KM KW
SQ SEQUENCE 808 AA; 89034 MW; 01BDF217F817D30 CRC64;

Query Match 79.5%; Score 35; DB 4; Length 808;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QYRAAGL 8
Db 665 KYRAAGL 672

RESULT 4
OY 09X0C1 PRELIMINARY; PRT; 80 AA.
AC 09X0C1;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE GLUTAREDOXIN.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999)
DR EMBL: AE001764; AAD36108.1; -
DR HSSP: P37687; IFOV.

```

DR TIGR; TM1031; -
 DR InterPro: IPR002109; Glutaredoxin.
 DR Pfam: PF00462; glutaredoxin; 1.
 DR PRINTS: PR00160; GLUTAREDOXIN.
 KM COMPLETE PROTEOME.
 SQ SEQUENCE 80 AA; 9127 MW; B1802B9D3460262F CRC64;

Query Match 77.3%; Score 34; DB 16; Length 80;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 : : : : :
 DB 22 EYFSLGLK 30

RESULT 5

ID 024020 PRELIMINARY; PRT; 187 AA.

AC 024020;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Cryptochrome (Fragment).

GN CYR2.

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RA Kojukisaoglu U., Schneider-Poetsch H.A.;

RT "The evolution of cryptochrome genes.";

RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ000695; CAA04247.1; -

DR InterPro: IPR005101; FAD_binding_7.

DR Pfam: PF03441; FAD_binding_7; 1.

DR NON_TER 1

FT NON_TER 1

FT NON_TER 1

SO SEQUENCE 187 AA; 21093 MW; 26D5459DD510E737 CRC64;

Query Match 77.3%; Score 34; DB 10; Length 187;
 Best Local Similarity 75.0%; Pred. No. 41;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
 : : : : :
 DB 147 YLRALGLK 154

RESULT 6

ID 09HDD6 PRELIMINARY; PRT; 361 AA.

AC 09HDD6;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Mannitol-1-phosphate dehydrogenase.

GN MPD1.

OS Cryptococcus neoformans var. neoformans.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;

OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.

OX NCBI_TaxID=40410;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H99;

RX MEDLINE-20476444; PubMed-11021946;

RA Suvarna K., Bartiss A., Wong B.;

RT "Mannitol-1-phosphate dehydrogenase from Cryptococcus neoformans is a

Microbiology 146:2705-2713(2000)."

DR EMBL: AF186474; MAG1013.1; -

DR EMBL: AF175685; AAC09209.1; -
 DR InterPro: IPR002085; adh_zn_family.
 DR Pfam: PF00107; adh_zinc; 1.
 SO SEQUENCE 361 AA; 37537 MW; F549926566CAF109 CRC64;

Query Match 77.3%; Score 34; DB 3; Length 361;
 Best Local Similarity 66.7%; Pred. No. 84;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 : : : : :
 DB 201 QYLRALGLK 209

RESULT 7

ID 09AGZ2 PRELIMINARY; PRT; 363 AA.

AC 09AGZ2;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Phosphoribosylaminoimidazole carboxylase.

GN PURK.

OS Ehrlichia canis.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;

OC Anaplasmataceae; Ehrlichia.

OX NCBI_TaxID=944;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-OKLAHOMA;

RX MEDLINE-98371112; PubMed-9705412;

RA Ohashi N., Unver A., Zhi N., Rikihisa Y.;

RT "Cloning and characterization of multigenes encoding the

immunodominant 30-kilodalton major outer membrane proteins of

Ehrlichia canis and application of the recombinant protein for

serodiagnosis.";

RL J. Clin. Microbiol. 36:2671-2680(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA STRAIN-OKLAHOMA;

RX MEDLINE-21153566; PubMed-11254561;

RA Ohashi N., Rikihisa Y., Unver A.;

RT "Analysis of transcriptionally Active Gene Clusters of Major Outer

Membrane Protein Multigene Family in Ehrlichia canis and E.

Chaffeensis.";

RL Infect. Immun. 69:2083-2091(2001).

DR EMBL: AF324792; AAK31311.1; -

DR HSSP; P09029; 1B6R.

DR InterPro: IPR003135; ATP-grasp.

DR Pfam: PF02222; ATP-grasp; 1.

DR TIGRfams: TIGR01161; purK; 1.

SO SEQUENCE 363 AA; 40744 MW; 5296634FCF56DE75 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 363;
 Best Local Similarity 66.7%; Pred. No. 84;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
 : : : : :
 DB 115 QYLRALGLK 123

RESULT 8

ID 09RXM0 PRELIMINARY; PRT; 406 AA.

AC 09RXM0;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Hypothetical protein DR0290.

GN DR0290.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE001890; AAF09871.1; -.
DR TIGR: DR0290; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 406 AA; 42901 MW; 6B40D3F1523F9FCB CRC64;

Query Match 77.3%; Score 34; DB 16; Length 406;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YLRALGL 8
|||
Db 33 YLRALGL 39

RESULT 9
O9LDJ5 PRELIMINARY; PRT; 635 AA.
AC O9LDJ5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cryptochrome 2.
OS Lycopodium obscurum (Lycopodiophyta).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20267445; PubMed=10809448;
RA Perrotta G., Ninn L., Flamma F., Weller J.L., Kendrick R.E.,
RA Nebuloso E., Giuliano G.;
RT "Tomato contains homologues of Arabidopsis cryptochromes 1 and 2.";
RL Plant Mol. Biol. 42:765-773(2000).
DR EMBL: AF130426; AAF72557.1; -.
DR EMBL: AF130425; AAF72556.1; -.
DR InterPro: IPR000474; DNA_photolyase.
DR InterPro: IPR002081; DNA_photolyase.
DR InterPro: IPR005101; DNA_photolyase.
DR Pfam: PF00875; DNA_photolyase.1.
DR Pfam: PF03441; FAD binding 7.1.
DR PRINTS: PR00147; DNAphotlyase.
DR PRODOM: PD004380; DNA_photolyase.1.
DR PROSITE: PS00394; DNA_PHOTOLYASES_1; 1.
SQ SEQUENCE 635 AA; 72249 MW; 937445E04DB86D6B CRC64;

Query Match 77.3%; Score 34; DB 10; Length 635;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 YLRALGL 9
|||
Db 284 YLRALGLR 291

RESULT 10
O9IBW6

ID O9IBW6 PRELIMINARY; PRT; 768 AA.
AC O9IBW6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE UL8.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GA;
RX MEDLINE=92237304; PubMed=1315048;
RA Jones D., Lee L., Liu J.L., Kung H.J., Tiliotson J.K.;
RT "Marek disease virus encodes a basic-leucine zipper gene resembling
RT the fos/jun oncogenes that is highly expressed in lymphoblastoid
RT tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4042-4046(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GA;
RX MEDLINE=20283955; PubMed=10823954;
RA Lee L.F., Wu P., Sul D., Ren D., Kamil J., Kung H.J., Witter R.L.;
RT "The complete unique long sequence and the overall genomic
RT organization of the GA strain of Marek's disease virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6091-6096(2000).
DR EMBL: AF147806; AAF66744.1; -.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR004996; Herpes_HEPA.
DR Pfam: PF03324; Herpes_HEPA.1.
DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN_1.
SQ SEQUENCE 768 AA; 86859 MW; E4CA706C4D185C88 CRC64;

Query Match 77.3%; Score 34; DB 12; Length 768;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYRALGL 9
|||
Db 671 OYRALGL 679

RESULT 11
O9EB08 PRELIMINARY; PRT; 769 AA.
AC O9EB08;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE UL8 DNA helicase-primase associated protein-like protein.
GN MDV020.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD5;
RX MEDLINE=20392152; PubMed=10933706;
RA Tulman E.R., Afonso C.L., Lu Z., Zaak L., Rock D.L., Kutish G.F.;
RT "The genome of a very virulent Marek's disease virus.";
RL J. Virol. 74:7980-7988(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MD5;
RA Tulman E.R., Afonso C.L., Lu Z., Zaak L., Rock D.L., Kutish G.F.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF243438; AAG14200.1; -.
DR InterPro: IPR004996; Herpes_HEPA.
DR Pfam: PF03324; Herpes_HEPA.1.
KW Helicase.

SO SEQUENCE 769 AA; 86907 MW; 28309BDBAF08333 CRC64;
 Query Match 77.3%; Score 34; DB 12; Length 769;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
 11:11:11
 Db 672 OYLRALGK 680

RESULT 12
 O9FKE9 PRELIMINARY; PRT; 1013 AA.
 ID O9FKE9;
 AC O9FKE9;
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative receptor-like serine/threonine kinase-partial protein
 DE (Fragment).
 DE F1M8.22.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federici N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Alatafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
 RA Buehler E., Chao O., Chin C., Chlou J., Choi E., Gonzalez A.,
 RA Hough B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
 RA Lenz C., Liu A., Liu S., Mukherjee N., Pham P., Sakano H., Shin P.,
 RA Tortum M., Vayenberg M., Yu G., Becker J., Theologis A., Davis R.W.,
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC008030; AAC10622.1;
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_Typ.
 DR InterPro: IPR002290; Ser_thr_Pkinase.
 DR InterPro: IPR004040; Sty_Pkinase.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR Pfam: PF00560; LRR; 7.
 DR Pfam: PF00069; Pkinase; 1.
 DR PRINTS: PR00019; LEURICHRP.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR SMART: SM00370; LRR; 5.
 DR SMART: SM00369; LRR_Typ; 2.
 DR SMART: SM00221; STYK; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR SMART: SM00219; TYKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Kinase; Receptor.
 FT NON_TER 1013 1013
 SO SEQUENCE 1013 AA; 112631 MW; 6A326BE4B2B10215 CRC64;

Query Match 77.3%; Score 34; DB 10; Length 1013;
 Best Local Similarity 77.8%; Pred. No. 2.6e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
 11:11:11
 Db 647 OYLRALGK 655

RESULT 13
 O8XK60 PRELIMINARY; PRT; 228 AA.
 AC O8XK60;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Dehydrobiotin synthetase.
 GN BI0D OR CPE1543.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL: AP003191; BAB1249.1;
 DR InterPro: IPR004472; BI0D.
 DR TIGRFAMs: TIGR00347; BI0D; 1.
 KW Complete proteome.
 SO SEQUENCE 228 AA; 25668 MW; 336974B1360E727D CRC64;

Query Match 75.0%; Score 33; DB 16; Length 228;
 Best Local Similarity 87.5%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 YLRALGK 9
 11:11:11
 Db 165 YLRALGK 172

RESULT 14
 O8R8F7 PRELIMINARY; PRT; 301 AA.
 ID O8R8F7;
 AC O8R8F7;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein TIE2045.
 GN TIE2045.
 OS Thermomicrobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermomicrobacteriales; Thermomicrobacteriaceae; Thermomicrobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao O., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.,
 RT "A complete sequence of T. tengcongensis genome."
 RT Genome Res. 12:689-700(2002).
 DR EMBL: AE013154; AAM25220.1;
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 301 AA; 35035 MW; 855F281C9CAD8297 CRC64;

Query Match 75.0%; Score 33; DB 16; Length 301;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGK 9
 11:11:11
 Db 67 EYRSIGLK 75

RESULT 15
 O8ZT63 PRELIMINARY; PRT; 314 AA.
 AC O8ZT63;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

```

DE 01-JUN-2002 (TRENBLER, 21, Last annotation update)
GN Hypothetical protein PAE3418.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Filz-Gibson S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009929; AAL64900.1; -
DR InterPro: IPR001279; B1actmase-1Like.
DR Pfam: PF00753; lactamase_B; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 314 AA; 34125 MW; 9B2EBB7107FD455 CRC64;

Query Match 75.0%; Score 33; DB 17; length 314;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 QYLRALGLK 9
: ||| : |||
Db 292 RYLRKIGLK 300

```

Search completed: July 11, 2003, 11:59:00
 Job time : 6.92931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 11.4761 Seconds
(Without alignments)
232.223 Million cell updates/sec

Title: US-10-091-442-6

Perfect score: 98

Sequence: 1 ADLSGIASGRLYSRMWCK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Listing first 45 summaries

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	93.9	20	15	AA857112
2	90	91.8	380	10	AA848379
3	90	91.8	380	20	AA108254
4	90	91.8	380	21	AA824142
5	90	91.8	380	22	AA833075
6	83	84.7	368	20	AA108256
7	83	84.7	368	21	AA824151
8	83	84.7	368	22	AA833077
9	83	84.7	380	20	AA108255
10	83	84.7	380	21	AA824150

11	83	84.7	380	22	AA64286
12	83	84.7	380	22	AA83076
13	50	51.0	615	22	AAU28187
14	50	51.0	615	22	AA83480
15	47	48.0	394	22	AA853280
16	47	48.0	395	22	AA847207
17	47	48.0	405	22	AA101601
18	47	48.0	617	22	ABG28280
19	46	46.9	327	22	AA890957
20	46	46.9	327	22	AA831515
21	46	46.9	618	22	ABG24021
22	45	45.9	307	22	AA632577
23	45	45.9	373	23	AB877879
24	45	45.9	387	23	AAU78358
25	45	45.9	406	23	AAU78357
26	44.5	45.4	434	12	AA813483
27	44	44.9	377	22	AA848382
28	44	44.9	453	23	AB848359
29	44	44.9	453	23	AB848359
30	43	43.9	135	22	AA722661
31	43	43.9	316	20	AA28664
32	43	43.9	391	20	AA41029
33	43	43.9	391	20	AA425926
34	43	43.9	391	21	AA811314
35	43	43.9	391	21	AA810953
36	43	43.9	391	21	AA810954
37	43	43.9	391	22	AA722651
38	43	43.9	391	23	AB874946
39	43	43.9	392	19	AAW48391
40	43	43.9	392	20	AAV04120
41	43	43.9	392	23	AAE14266
42	43	43.9	400	20	AAV41030
43	43	43.9	400	21	AA811315
44	43	43.9	400	22	AA722652
45	43	43.9	400	23	AB874947

ALIGNMENTS

RESULT 1	AA857112	standard; peptide: 20 AA.
XX	AA857112:	
AC	AA857112:	
XX	16-AUG-1994 (first entry)	
XX		
DE	Human megakaryocyte differentiation factor peptide 6.	
XX		
KW	Human megakaryocyte differentiation factor; MDF; thrombopoietin;	
KW	hematopoietic stimulating factor; thrombocytopoietin; platelet;	
KW	bone marrow transplantation; cancer chemotherapy.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 18	/note= "not determined"
XX		
PN	EP583884-A.	
PD	23-FEB-1994.	
XX		
FE	19-JUL-1993;	93EP-0305654.
XX		
PR	17-JUL-1992;	92JP-0212305.
XX	04-MAR-1993;	93JP-0067339.
XX		
PA	(SUNR) SUNTORY LTD.	
PA	(TSUO/) TSUDIMOTO M.	
PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;	

Rat megilin protein
Rat megilin protein
Novel human secret
Human protein sequ
Human polypeptide
Human NOVI protein
Protein encoded by
Novel human diago
Canine GnRH recept
A dog gonadotropin
Novel human diago
C glutamicum prote
Amino acid sequenc
Modified human pro
Human protein C in
P. denitrificans CO
I. ricinus salivar
Modified human pro
Listeria monocytog
Human headpin (for
Human serine prote
Human lung tumor a
Human huplin prote
Human lung cancer-
Human serine prote
Human serine prote
Human headpin (for
Human lung tumour
Human lung tumour
Hom sapiens pancr
Pancreas derived p
Human Pancreas-der
Human lung tumor a
Human lung cancer-
Human headpin (hea
Human lung tumour

PI Tsurumoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
 XX WPI: 1994-058782/08.
 XX New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 decrease in platelets
 XX
 PS Claim 1; Page 20; 47pp: English.
 XX
 CC Human MDF (see AAR48379) can be isolated from a culture of human
 CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
 CC stimulates differentiation of megakaryocytes from myeloid cells
 CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
 CC making it useful for treatment of diseases involving a decrease
 CC in platelet number (esp. thrombocytopaenia) such as occurs in bone
 CC marrow transplantation and in chemotherapy. MDF has mol.wt.
 CC 55-57kD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
 CC contains an amino acid sequence comprising at least one of the
 CC sequences AAR57107-R57115.
 XX
 SQ Sequence 20 AA;
 Query Match 93.9%; Score 92; DB 15; Length 20;
 Best Local Similarity 95.0%; Pred. No. 5e-08;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ADLSGIASGGRLYISRMKG 20
 DB 1 ADLSGIASGGRLYISRMKG 20
 RESULT 2
 AAR48379
 ID AAR48379 standard; Protein: 380 AA.
 XX
 AC AAR48379;
 XX
 DT 16-AUG-1994 (first entry)
 XX
 DE Human megakaryocyte differentiation factor.
 XX
 KM Human megakaryocyte differentiation factor; MDF; thrombopoietin;
 KM haematopoietic stimulating factor; thrombocytopaenia; platelet;
 KM bone marrow transplantation; cancer chemotherapy.
 XX
 OS Homo sapiens.
 XX
 PN EP583884-A.
 XX
 PD 23-FEB-1994.
 XX
 PF 19-JUL-1993; 93EP-0305654.
 XX
 PR 17-JUL-1992; 92JP-0212305.
 PR 04-MAR-1993; 93JP-0067339.
 XX
 PA (SUNR) SUNTORY LTD.
 PA (TSUJ) TSUJIMOTO M.
 XX
 PI Iehida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
 PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
 XX
 DR WPI: 1994-058782/08.
 DR N-PSDB; AAO56670.
 XX
 PT New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 decrease in platelets
 XX
 PS Claim 7; Page 30-32; 47pp: English.
 XX
 CC Human MDF can be isolated from a culture of human epidermoid

CC carcinoma A431 cells in protein-free medium. The MDF stimulates
 CC differentiation of megakaryocytes from myeloid cells in the presence
 CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
 CC for treatment of diseases involving a decrease in platelet number
 CC (esp. thrombocytopaenia) such as occurs in bone marrow
 CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by
 CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
 XX
 SQ Sequence 380 AA;
 Query Match 91.8%; Score 90; DB 15; Length 380;
 Best Local Similarity 95.0%; Pred. No. 2.7e-06;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ADLSGIASGGRLYISRMKG 20
 DB 305 ADLSGIASGGRLYISRMKG 324
 RESULT 3
 AAY08254
 ID AAY08254 standard; Protein: 380 AA.
 XX
 AC AAY08254;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Human megalin protein.
 XX
 KM Megalin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 KM human; rat; murine.
 XX
 OS Homo sapiens.
 XX
 PN WO915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98WO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO) KUROKAWA K.
 PA (MIYA) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 1999-276983/23.
 DR N-PSDB; AAX56712.
 XX
 PT Megalin protein expressed specifically in mesangial cells
 XX
 PS Claim 1; Page 62-64; 100pp; Japanese.
 XX
 CC This invention describes the isolation of novel megalin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as Iga nephropathy.
 XX
 SQ Sequence 380 AA;
 Query Match 91.8%; Score 90; DB 20; Length 380;
 Best Local Similarity 95.0%; Pred. No. 2.7e-06;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ADLSGIASGGRLYISRMKG 20
 DB 305 ADLSGIASGGRLYISRMKG 324
 RESULT 4
 AAB24142

ID AAB24142 standard; Protein: 380 AA.
 XX
 AC AAB24142;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Human megalin protein sequence SEQ ID NO:2.
 XX
 KW Megalin: mesangial-predominant gene; serpin regulated; nephropathy;
 KW Iga; immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample: blood; urine.
 XX
 OS Homo sapiens.
 XX
 PN W0200057189-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 17-MAR-2000; 2000MO-JP01646.
 XX
 PR 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KURO/) KUROKAWA K.
 PA (FUSO) FUSO PHARM IND LTD.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2000-611642/58.
 DR N-PSDB: AAA99294.
 XX
 PT Evaluating renal function comprises assaying megalin protein in
 PT biological sample -
 XX
 Example 2; Page 66-69; 93pp; Japanese.
 XX
 PS The present invention describes a method for evaluating renal function.
 PS The method comprises assaying megalin protein in biological sample. Also
 CC described are: (1) use of an anti-megalin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megalin protein comprising:
 CC (a) anti-megalin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megalin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the human megalin
 CC protein, which is given in the exemplification of the present invention.
 XX
 SQ Sequence 380 AA;
 XX
 QY Query Match 91.8%; Score 90; DB 21; Length 380;
 QY Best Local Similarity 95.0%; Pred. No. 2,7e-06;
 QY Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DB 1 ADLSGIASGRLYISRMWK 20
 DB 305 ADLSGIASGRLYISRMWK 324
 XX
 RESULT 5
 ID AAB83075 standard; Protein: 380 AA.
 XX
 AC AAB83075;
 XX
 DT 10-JUL-2001 (first entry)
 XX
 DE Human megalin protein.
 XX
 KW Human; megalin; mesangial cell proliferative nephritis; nephrotropic;
 KW transgenic mouse; glomerular disease; animal model; drug screening.
 XX

OS Homo sapiens.
 XX
 PN W0200124628-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 06-OCT-2000; 2000MO-JP06988.
 XX
 PR 06-OCT-1999; 99JP-0285736.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2001-300136/31.
 DR N-PSDB: AAF82438.
 XX
 PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments -
 XX
 PS Example 4; Page 44-46; 62pp; Japanese.
 XX
 PS The present sequence is human megalin. The human megalin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 XX
 SQ Sequence 380 AA;
 XX
 QY Query Match 91.8%; Score 90; DB 22; Length 380;
 QY Best Local Similarity 95.0%; Pred. No. 2,7e-06;
 QY Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DB 1 ADLSGIASGRLYISRMWK 20
 DB 305 ADLSGIASGRLYISRMWK 324
 XX
 RESULT 6
 ID AAY08256 standard; Protein: 368 AA.
 XX
 AC AAY08256;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Mouse megalin protein.
 XX
 KW Megalin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 KW human; rat; murine.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FH Protein 1..368
 FT /note="partial sequence"
 FT XX
 PN W09915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98MO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.


```

RESULT 9
AAV08255
ID AAV08255 standard; Protein: 380 AA.
XX
AC AAV08255;
XX
DT 14-JUL-1999 (first entry)
XX
DE Rat megalin protein.
XX
KW Megalin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
KM human; rat; murine.
XX
OS Rattus rattus.
XX
PN WO915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
XX
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 1999-276983/23.
DR N-PSDB: AAX56712.
XX
PT Megalin protein expressed specifically in mesangial cells
PS Claim 1; Page 69-72; 100pp: Japanese.
XX
CC This invention describes the isolation of novel megalin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as Iga nephropathy.
XX
SQ Sequence 380 AA;
OY
Query Match 84.7%; Score 83; DB 20; Length 380;
Best Local Similarity 80.0%; Pred. No. 3.6e-05;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 1 ADLGGIAGGRLYISRMGK 20
305 ADLGGIAGGRLYISRMGK 324

RESULT 10
AAB24150
ID AAB24150 standard; Protein: 380 AA.
XX
AC AAB24150;
XX
DT 30-JAN-2001 (first entry)
XX
DE Rat megalin protein sequence SEQ ID NO:19.
XX
KW Megalin; mesangium-predominant gene; serpin regulated; nephropathy;
KM Iga; immunoglobulin A; detection; renal function; renal disorder;
diagnosis; biological sample; blood; urine.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT Misc-difference 51 /note= "unspecified"
FT Misc-difference 94
FT /note= "unspecified"

```

```

XX
XX WO200057189-A1.
XX
XX 28-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-JP01646.
XX
XX 19-MAR-1999; 99JP-0075305.
XX
XX 28-OCT-1999; 99JP-0306623.
XX
PA (KURO/) KUROKAWA K.
PA (FUSO ) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
XX Miyata T;
XX
XX WPI: 2000-611642/58.
XX
XX N-PSDB: AAC55238.
XX
XX
XX Evaluating renal function comprises assaying megalin protein in
XX biological sample
XX
XX Example 2; Page 81-84; 93pp: Japanese.
XX
XX
XX The present invention describes a method for evaluating renal function.
XX The method comprises assaying megalin protein in biological sample. Also
XX described are: (1) use of a anti-megalin protein antibody for diagnosing
XX renal function; and (2) a kit for detecting megalin protein comprising:
XX (a) anti-megalin protein antibody attached to solid magnetic particles;
XX (b) direct or indirect fixing for the antibody to the particles; and
XX (c) a magnet. The process is useful for evaluating renal function and
XX diagnosing renal disorders by assaying megalin protein in biological
XX samples (preferably urine or blood). The process is reproducible and
XX gives accurate results. The present sequence represents the rat megalin
XX protein, which is given in the exemplification of the present invention.
XX
XX
XX Sequence 380 AA;
OY
Query Match 84.7%; Score 83; DB 21; Length 380;
Best Local Similarity 80.0%; Pred. No. 3.6e-05;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 1 ADLGGIAGGRLYISRMGK 20
305 ADLGGIAGGRLYISRMGK 324

RESULT 11
AAG64286
ID AAG64286 standard; Protein: 380 AA.
XX
XX AAG64286;
XX
XX 21-SEP-2001 (first entry)
XX
XX Rat megalin protein.
XX
XX Rat; megalin; renal mesangial cell; mesangium proliferative nephritis.
XX
XX Rattus norvegicus.
XX
XX WO200148019-A1.
XX
XX 05-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-JP09251.
XX
XX 28-DEC-1999; 99JP-0373677.
XX
XX (KURO/) KUROKAWA K.
XX (MIYA/) MIYATA T.
XX
XX Miyata T;

```

XX WPI: 2001-425651/45.
DR N-PSDB: AAH48181.
XX
PT New antibody recognizing a partial sequence of rat mesgin protein for
PT diagnosis of mesangium proliferative nephritis
XX
PS Disclosure: Page 54-56; 63pp; Japanese.
XX
CC The present invention relates to a novel antibody which recognises a
CC peptide consisting of residues 341-354 of rat mesgin protein. The present
CC sequence is the protein sequence for rat mesgin, which was used in the
CC present invention. Mesgin is highly expressed in renal mesangial cells
CC and its level is elevated in mesangium proliferative nephritis. Assay of
CC the serum or urine level using the antibody is therefore indicative of
CC this type of disorder.
XX
SQ Sequence 380 AA:
Query Match 84.7%; Score 83; DB 22; Length 380;
Best Local Similarity 80.0%; Pred. No. 3.6e-05;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ADLSGIASGRLYISRMWGK 20
Db 305 ADLSGIASGRLYISKLMHK 324
|||||||
RESULT 12.
AAB83076
ID AAB83076 standard; Protein; 380 AA.
XX
AC AAB83076:
XX
DT 10-JUL-2001 (first entry)
XX
DE Rat mesgin protein.
XX
DE Rat mesgin protein.
XX
KW Rat; mesgin; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening.
XX
OS Rattus norvegicus.
XX
PN W0200124628-A1.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000MO-JP06988.
XX
PR 06-OCT-1999; 99JP-0285736.
XX
PA (KUROO) KUROKAWA K.
PA (MIYA) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 2001-300136/31.
DR N-PSDB: AAF82439.
XX
PT Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments
XX
PS Disclosure: Page 48-50; 63pp; Japanese.
XX
CC The present sequence is rat mesgin. The human mesgin coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.

XX SQ Sequence 380 AA:
Query Match 84.7%; Score 83; DB 22; Length 380;
Best Local Similarity 80.0%; Pred. No. 3.6e-05;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ADLSGIASGRLYISRMWGK 20
Db 305 ADLSGIASGRLYISKLMHK 324
|||||||
RESULT 13
AAU28187
ID AAU28187 standard; Protein; 615 AA.
XX
AC AAU28187:
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secretory protein, Seq ID No 356.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen.
XX
OS Homo sapiens.
XX
PN W020016689-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001MO-US04942.
XX
PR 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HXSE-) HXSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR WPI: 2001-589934/66.
DR N-PSDB: AAS45087.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders
XX
PS Example 5; SEQ ID No 356; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing hematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,

CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (1) is involved in chemotactic or chemokinetic
 CC activity, regulation of hematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (1) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (1) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.

XX Sequence 615 AA;

SO Query Match. 51.0%; Score 50; DB 22; Length 615;

Best Local Similarity 83.3%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 SGASGRLYIS 15

DB 454 AGASGRLYIS 465

RESULT 14

AAB93480 ID AAB93480 standard; Protein: 615 AA.

XX AAB93480;

XX 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12768.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EP1074617-A2.

PN 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 615 AA;

SO Query Match. 51.0%; Score 50; DB 22; Length 615;

Best Local Similarity 83.3%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 SGASGRLYIS 15

DB 454 AGASGRLYIS 465

RESULT 15

ABB53280 ID ABB53280 standard; Protein: 394 AA.

XX ABB53280;

XX 12-FEB-2002 (first entry)

DE Human polypeptide #20.

KW Human; nootropic; neuroprotective; anticonvulsant; antidepressant;

KW neuroleptic; tranquiliser; antiarrhythmic; cardiant; antidiabetic;

KW antiinflammatory; antihypertensive; hepatotropic; virucide; antidiabetic;

KW nephrotropic; anorectic; cytotatic; vaccine; neurological disease;

KW cardiovascular disease; respiratory disease; liver disease;

KW renal disease; skeletal muscle disease; gastrointestinal disease;

KW placental disease; testicular cancer; male fertility; pancreatic disease.

XX Homo sapiens.

XX WO200181363-A1.

XX 01-NOV-2001.

XX 26-APR-2001; 2001WO-US13360.

XX 27-APR-2000; 2000US-199963P.

XX 11-MAY-2000; 2000US-203336P.

XX 25-MAY-2000; 2000US-207087P.

XX 26-MAY-2000; 2000US-207546P.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;

XX WPI: 2002-041392/05.

XX N-PSDB; ABA90345.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 3.57588 Seconds

(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-6

Sequence: 1 ADLSGIASGRLYSRMWK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents_Aa:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	98	100.0	20	2	US-08-611-977-6	Sequence 6, Appl1	
2	92	93.9	20	2	US-08-472-659-6	Sequence 6, Appl1	
3	92	93.9	20	2	US-08-474-661-6	Sequence 6, Appl1	
4	90	91.8	380	2	US-08-472-659-34	Sequence 34, Appl1	
5	90	91.8	380	2	US-08-474-661-34	Sequence 34, Appl1	
6	90	91.8	380	2	US-08-611-977-34	Sequence 34, Appl1	
7	46	46.9	327	4	US-09-583-492-10	Sequence 10, Appl1	
8	43	43.9	375	1	US-08-121-714-8	Sequence 8, Appl1	
9	43	43.9	375	1	US-08-477-108A-8	Sequence 8, Appl1	
10	43	43.9	375	2	US-08-477-112-8	Sequence 8, Appl1	
11	43	43.9	375	5	PCT-US93-08322-8	Sequence 8, Appl1	
12	43	43.9	391	4	US-09-123-912-110	Sequence 110, Appl1	
13	43	43.9	391	4	US-09-643-597-110	Sequence 110, Appl1	
14	43	43.9	392	4	US-09-026-408-2	Sequence 2, Appl1	
15	43	43.9	400	4	US-09-123-912-112	Sequence 112, Appl1	
16	43	43.9	400	4	US-09-643-597-112	Sequence 112, Appl1	
17	43	43.9	405	4	US-09-026-408-13	Sequence 13, Appl1	
18	43	43.9	406	1	US-08-434-881-2	Sequence 2, Appl1	
19	43	43.9	406	3	US-08-977-771-2	Sequence 2, Appl1	
20	43	43.9	406	4	US-09-361-773-2	Sequence 2, Appl1	
21	42	42.9	379	1	US-08-121-714-4	Sequence 4, Appl1	
22	42	42.9	379	1	US-08-477-108A-4	Sequence 4, Appl1	
23	42	42.9	379	2	US-08-477-112-4	Sequence 4, Appl1	
24	42	42.9	379	5	PCT-US93-08322-4	Sequence 4, Appl1	
25	41.5	42.3	331	4	US-09-199-637A-333	Sequence 333, Appl1	
26	41	41.8	245	4	US-09-342-084-4	Sequence 4, Appl1	
27	39	39.8	65	4	US-09-579-236-5	Sequence 5, Appl1	

28	39	39.8	125	4	US-09-579-236-2	Sequence 2, Appl1
29	39	39.8	328	1	US-08-423-691-2	Sequence 2, Appl1
30	39	39.8	328	1	US-08-080-386-4	Sequence 4, Appl1
31	39	39.8	328	2	US-08-390-000A-4	Sequence 4, Appl1
32	39	39.8	328	3	US-08-867-260-2	Sequence 2, Appl1
33	39	39.8	465	3	US-09-501-192-9	Sequence 9, Appl1
34	39	39.8	465	3	US-08-948-997-6	Sequence 6, Appl1
35	39	39.8	465	4	US-09-348-817A-6	Sequence 6, Appl1
36	39	39.8	489	2	US-08-752-307B-3	Sequence 3, Appl1
37	39	39.8	489	4	US-09-330-317B-18	Sequence 18, Appl
38	39	39.8	489	4	US-09-707-802-3	Sequence 3, Appl1
39	39	39.8	489	4	US-09-991-326-3	Sequence 3, Appl1
40	39	39.8	489	4	US-09-808-589A-18	Sequence 18, Appl
41	39	39.8	506	4	US-09-360-237-59	Sequence 59, Appl
42	39	39.8	530	2	US-08-752-307B-2	Sequence 2, Appl1
43	39	39.8	530	4	US-09-707-802-2	Sequence 2, Appl1
44	39	39.8	530	4	US-09-991-326-2	Sequence 2, Appl1
45	39	39.8	625	2	US-08-532-547-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-08-611-977-6
Sequence 6, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURU, Kenju
APPLICANT: ISHIDA, No. 5972886uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm R.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-6

Query Match 100.0%; Score 98; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXGK 20
Db 1 ADLSGIASGRLYISRMXGK 20

RESULT 2

US-08-472-659-6
Sequence 6, Application US/08472659
Patent No. 5831030

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IMASA, Fuyuki

APPLICANT: TSUROJOKA, No. 5831030uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5831030unhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAICHI, Kozo

APPLICANT: YAMAGUCHI, No. 5831030oml

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,659

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET INFORMATION: 001560-248

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-472-659-6

Query Match 93.9%; Score 92; DB 2; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.1e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXGK 20
Db 1 ADLSGIASGRLYISRMXGK 20

RESULT 3

US-08-474-661-6
Sequence 6, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IMASA, Fuyuki

APPLICANT: TSUROJOKA, No. 5874253uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5874253unhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAICHI, Kozo

APPLICANT: YAMAGUCHI, No. 5874253oml

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: George Mason Bldg., Washington & Prince Sts.

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,661

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: REA, TERESA STANER

REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET INFORMATION: 001560-204

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-474-661-6

Query Match 93.9%; Score 92; DB 2; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.1e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXGK 20
Db 1 ADLSGIASGRLYISRMXGK 20

RESULT 4
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472.659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 91.8%; Score 90; DB 2; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.6e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMWK 20
Db 305 ADLSGIASGRLYISRMWK 324

RESULT 5
US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki

APPLICANT: TSUROOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474.661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 91.8%; Score 90; DB 2; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.6e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMWK 20
Db 305 ADLSGIASGRLYISRMWK 324

RESULT 6
US-08-611-977-34
Sequence 34, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-611-977-34

Query Match 91.8%; Score 90; DB 2; Length 380;
Best Local Similarity 95.0%; Pred. No. 1,6e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMK 20
Db 305 ADLSGASGRLYISRMHK 324

RESULT 7
US-09-583-492-10
Sequence 10, Application US/09583492
Patent No. 6407222
GENERAL INFORMATION:
APPLICANT: Cul, Jisong
APPLICANT: Lo, Jane-Ling
APPLICANT: Mount, George R.
TITLE OF INVENTION: DOG GONADOTROPIN RELEASING HORMONE
FILE REFERENCE: 20140
CURRENT APPLICATION NUMBER: US/09/583,492
CURRENT FILING DATE: 2000-05-31
EARLIER APPLICATION NUMBER: 60/138,448
EARLIER FILING DATE: 2000-06-10
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 327
TYPE: PRT
ORGANISM: canis
US-09-583-492-10

Query Match 46.9%; Score 46; DB 4; Length 327;

Best Local Similarity 62.5%; Pred. No. 4.2;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LSGIASGRLYISRM 18
Db 165 LSGIFAGPOLYIFPMI 180

RESULT 8
US-08-121-714-8
Sequence 8, Application US/08121714
Patent No. 5470970
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Mordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,714
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 375
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-121-714-8

Query Match 43.9%; Score 43; DB 1; Length 375;
Best Local Similarity 40.0%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMK 20
Db 302 ADLSGMSGARDLFVSKTIHK 321

RESULT 9
US-08-477-1084-8
Sequence 8, Application US/084771084
Patent No. 5801001
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Zou, Zhigang
APPLICANT: Anisowicz, Anthony

;; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
;; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 225 Franklin Street
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: U.S.A.
;; ZIP: 02110-2804
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 502 or 55SX
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,108A
;; FILING DATE: June 7, 1995
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/121,714
;; FILING DATE: 09/01/93
;; APPLICATION NUMBER: 07/938,823
;; FILING DATE: 09/01/92
;; APPLICATION NUMBER: 07/844,296
;; FILING DATE: 02/28/92
;; APPLICATION NUMBER: 07/662,216
;; FILING DATE: 02/28/91
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fraser, Janis K.
;; REGISTRATION NUMBER: 34,819
;; REFERENCE/DOCKET NUMBER: 06570/002002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 542-5070
;; TELEFAX: (617) 542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 375
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;;
;; US-08-477-108A-8
;;
Query Match 43.9%; Score 43; DB 1; Length 375;
Best Local Similarity 40.0%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
;;
Qy 1 ADLSCIASGRLYISRMGK 20
DB 302 ADLSCIASGRLYISRMGK 321
;;
RESULT 10
US-08-477-112-8
;; Sequence 8, Application US/08477112
;; Patent No. 5905023
;; GENERAL INFORMATION:
;; APPLICANT: Sager, Ruth
;; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
;; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 225 Franklin Street
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: U.S.A.
;; ZIP: 02110-2804
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 502 or 55SX
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version 5.0)

;; SOFTWARE: WordPerfect (Version 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,112
;; FILING DATE:
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/121,714
;; FILING DATE: 09/01/93
;; APPLICATION NUMBER: 07/938,823
;; FILING DATE: 09/01/92
;; APPLICATION NUMBER: 07/844,296
;; FILING DATE: 02/28/92
;; APPLICATION NUMBER: 07/662,216
;; FILING DATE: 02/28/91
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fraser, Janis K.
;; REGISTRATION NUMBER: 34,819
;; REFERENCE/DOCKET NUMBER: 06570/002003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 542-5070
;; TELEFAX: (617) 542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 375
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;;
;; US-08-477-112-8
;;
Query Match 43.9%; Score 43; DB 2; Length 375;
Best Local Similarity 40.0%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
;;
Qy 1 ADLSCIASGRLYISRMGK 20
DB 302 ADLSCIASGRLYISRMGK 321
;;
RESULT 11
PCT-US93-08322-8
;; Sequence 8, Application PC/TUS9308322
;; GENERAL INFORMATION:
;; APPLICANT: Sager, Ruth
;; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson
;; STREET: 225 Franklin Street
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: U.S.A.
;; ZIP: 02110-2804
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 502 or 55SX
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/08322
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/938,823
;; FILING DATE: 09/01/92
;; APPLICATION NUMBER: 07/844,296
;; FILING DATE: 02/28/92
;; APPLICATION NUMBER: 07/662,216
;; FILING DATE: 02/28/91
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fraser, Janis K.
;; REGISTRATION NUMBER: 34,819
;; REFERENCE/DOCKET NUMBER: 00530/072001

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-08322-8

Query Match      43.9%; Score 43; DB 5; Length 375;
Best Local Similarity 40.0%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 ADLSGASGRLYISRMKG 20
Db      302 ADLSGMSGARDLFSKTIHK 321

RESULT 12
US-09-123-912-110
; Sequence 110, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455c1
; CURRENT APPLICATION NUMBER: US/09/123.912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040.802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 110
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-123-912-110

Query Match      43.9%; Score 43; DB 4; Length 391;
Best Local Similarity 44.4%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      1 ADLSGASGRLYISRM 18
Db      314 ADYSGMSSGSLYAOKRL 331

RESULT 13
US-09-643-597-110
; Sequence 110, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455c11
; CURRENT APPLICATION NUMBER: US/09/643.597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-643-597-110

Query Match      43.9%; Score 43; DB 4; Length 391;
Best Local Similarity 44.4%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      1 ADLSGASGRLYISRM 18
Db      314 ADYSGMSSGSLYAOKRL 331

RESULT 14
US-09-026-408-2
; Sequence 2, Application US/09026408
; Patent No. 630338
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026.408
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/934.011
; FILING DATE: 13-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024.056
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0300002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-026-408-2

Query Match      43.9%; Score 43; DB 4; Length 392;
Best Local Similarity 42.1%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 DLSGASGRLYISRMKG 20
Db      316 DLSGTDSSEVYVSOYOK 334

RESULT 15
US-09-123-912-112
```

```

; Sequence 112, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455c1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 112
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-123-912-112

```

```

Query Match      43.9%; Score 43; DB 4; Length 400;
Best Local Similarity 44.4%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
OY      1 ADLSGIASGGRLYISRM 18
        |||||:|  ||  :
Db      323 ADYSGMSSGSLYAKFL 340

```

Search completed: July 11, 2003, 12:02:09
Job time : 4.57588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 5.86279 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-6
Perfect score: 98
Sequence: 1 ADLSGASGGRLYSRMGK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCIT_NEM_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCITUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	20	9 US-10-091-442-6	Sequence 6, Appl1
2	98	100.0	20	10 US-09-140-719-6	Sequence 6, Appl1
3	90	91.8	380	9 US-10-091-442-34	Sequence 34, Appl1
4	90	91.8	380	10 US-09-140-719-34	Sequence 34, Appl1
5	47	48.0	395	9 US-09-823-187-2	Sequence 2, Appl1
6	45	45.9	307	9 US-09-738-626-6331	Sequence 6331, App
7	45	45.9	379	9 US-10-097-340-139	Sequence 139, App
8	44	44.9	377	9 US-10-165-605A-27	Sequence 27, App
9	44	44.9	377	10 US-09-910-430-37	Sequence 27, Appl1
10	43	43.9	391	9 US-10-157-031-355	Sequence 355, App
11	43	43.9	391	9 US-09-466-396A-110	Sequence 110, App
12	43	43.9	391	10 US-09-735-705-110	Sequence 110, App
13	43	43.9	391	10 US-09-850-716A-110	Sequence 110, App
14	43	43.9	391	10 US-09-897-778-110	Sequence 110, App
15	43	43.9	392	10 US-09-903-684-2	Sequence 2, Appl1
16	43	43.9	400	9 US-09-466-396A-112	Sequence 112, App
17	43	43.9	400	10 US-09-735-705-112	Sequence 112, App
18	43	43.9	400	10 US-09-850-716A-112	Sequence 112, App
19	43	43.9	400	10 US-09-897-778-112	Sequence 112, App

20	43	43.9	405	10 US-09-902-684-13	Sequence 13, Appl1
21	42	42.9	437	10 US-09-925-301-1200	Sequence 1200, App
22	41.5	42.3	331	9 US-09-975-719-333	Sequence 333, App
23	41	41.8	143	9 US-10-091-504-679	Sequence 679, App
24	41	41.8	143	10 US-09-764-869-679	Sequence 679, App
25	41	41.8	217	10 US-09-729-674-134	Sequence 134, App
26	41	41.8	340	10 US-09-764-898-169	Sequence 169, App
27	41	41.8	379	9 US-10-012-896-946	Sequence 946, App
28	41	41.8	392	9 US-09-895-814-946	Sequence 946, App
29	41	41.8	392	9 US-10-198-070-62	Sequence 62, Appl1
30	41	41.8	392	9 US-10-198-070-67	Sequence 67, Appl1
31	41	41.8	392	9 US-10-198-070-70	Sequence 70, Appl1
32	41	41.8	402	9 US-10-135-629-4	Sequence 4, Appl1
33	41	41.8	413	9 US-10-223-857-4	Sequence 4, Appl1
34	41	41.8	423	9 US-09-993-180-3	Sequence 3, Appl1
35	41	41.8	423	9 US-10-135-629-3	Sequence 3, Appl1
36	41	41.8	423	9 US-09-919-039-19	Sequence 19, Appl1
37	41	41.8	423	9 US-09-978-418-14	Sequence 14, Appl1
38	41	41.8	433	9 US-10-205-823-361	Sequence 361, App
39	41	41.8	433	9 US-10-177-293-421	Sequence 421, App
40	41	41.8	617	9 US-10-012-896-947	Sequence 947, App
41	41	41.8	617	9 US-09-895-814-947	Sequence 947, App
42	40	40.8	91	10 US-09-925-302-886	Sequence 886, App
43	39	39.8	65	10 US-09-754-016-5	Sequence 5, Appl1
44	39	39.8	125	10 US-09-754-016-2	Sequence 2, Appl1
45	39	39.8	328	9 US-10-225-567A-136	Sequence 136, App

ALIGNMENTS

RESULT 1

US-10-091-442-6
Sequence 6, Application US/10091442

Patent No. US20020164711A1
GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

INVENTOR: IWASA, Fuyuki

ATTORNEY: TSURUOKA, No. US20020164711A1uo

INVENTOR: NAKAZATO, Hiroshi

INVENTOR: MIURA, Kenju

INVENTOR: ISHIDA, No. US20020164711A1uhfro

INVENTOR: KURIHARA, Tatsuya

INVENTOR: YAMACHI, Koza

INVENTOR: YAMAGUCHI, No. US20020164711A1oml

TITLE OF INVENTION: MEKANARYCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-091-442-6

Query Match 100.0%; Score 98; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSGIASGGRLYISRMGK 20
DB 1 ADLSGIASGGRLYISRMGK 20

RESULT 2
US-09-140-719-6
Sequence 6, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-6

Query Match 100.0%; Score 98; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADLSGIASGGRLYISRMGK 20
DB 1 ADLSGIASGGRLYISRMGK 20

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-MAR-2002
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids,
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 91.8%; Score 90; DB 9; Length 380;
Best Local Similarity 95.0%; Pred. NO. 2e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMKG 20
DB 305 ADLSGASGRLYISRMKG 324

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUBOUOKA, No. US20010026931A1uo
APPLICANT: NAKAZARO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 91.8%; Score 90; DB 10; Length 380;
Best Local Similarity 95.0%; Pred. NO. 2e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMKG 20
DB 305 ADLSGASGRLYISRMKG 324

RESULT 5
US-09-823-187-2
Sequence 2, Application US/09823187
Publication No. US20030096952A1
GENERAL INFORMATION:
APPLICANT: Burgess, Catherine
APPLICANT: Gusev, Vladimir Y
APPLICANT: Liu, Xiaohong
APPLICANT: Majumder, Kumud
APPLICANT: Padigaru, Muralidhar
APPLICANT: Pattnayan, Meera
APPLICANT: Shimkets, Richard A
APPLICANT: Spaderna, Steven K
APPLICANT: Spytek, Kimberly
APPLICANT: Taupier, Raymond J
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-745
CURRENT APPLICATION NUMBER: US/09/823,187
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/193,339
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/193,205
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/195,343
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: 60/195,088
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,005
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,792
PRIOR FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: 60/196,556
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: 60/197,081
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 60/197,525
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/197,087
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 103
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 395
TYPE: PRT
ORGANISM: Homo sapiens
US-09-823-187-2

Query Match 48.0%; Score 47; DB 9; Length 395;
Best Local Similarity 45.0%; Pred. NO. 14;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMKG 20
DB 318 ADLTGISPSPNLYLSKITHK 337

RESULT 6
US-09-738-626-6331
Sequence 6331, Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6331
LENGTH: 307
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6331

Query Match 45.9%; Score 45; DB 9; Length 307;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 LSGIASGRLYISRMKG 20
DB 215 LSGIYGVMLAVARVGE 232

RESULT 7

US-10-097-340-139
Sequence 139, Application US/10097340
Publication No. US20030087250A1
GENERAL INFORMATION:
APPLICANT: JOHN MONAHAN
APPLICANT: Manjula GANNAVAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732

PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 139
LENGTH: 379
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-139

Query Match 45.9%; Score 45; DB 9; Length 379;
Best Local Similarity 45.0%; Pred. No. 28;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMKG 20
DB 304 ADLSGISHNSNTQVSEVHK 323

RESULT 8

US-10-165-605A-27
Sequence 27, Application US/10165605A
Publication No. US20030086937A1
GENERAL INFORMATION:
APPLICANT: Godfroid, Edmond
APPLICANT: Bollen, Alex
APPLICANT: Leboulle, Gerard
TITLE OF INVENTION: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF
TITLE OF INVENTION: PROTEINS, EXPRESSED IN THE IXODS RICINUS SALIVARY
FILE REFERENCE: VANM229.001CP2
CURRENT APPLICATION NUMBER: US/10/165,605A
CURRENT FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 09/910,430
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: PCT/BE00/00061
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: GB9913425.6
PRIOR FILING DATE: 1998-06-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 377
TYPE: PRT
ORGANISM: Ixodes ricinus
US-10-165-605A-27

Query Match 44.9%; Score 44; DB 9; Length 377;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 ADLSGASGRLYISRMKG 20
DB 298 ADLSGASAGNLCAVDVHK 317

RESULT 9

US-09-910-430-27
Sequence 27, Application US/09910430
Patent No. US20020127235A1
GENERAL INFORMATION:
APPLICANT: Godfroid, Edmond
APPLICANT: Bollen, Alex
APPLICANT: Leboulle, Gerard
TITLE OF INVENTION: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF
TITLE OF INVENTION: PROTEINS, EXPRESSED IN THE IXODS RICINUS SALIVARY
FILE REFERENCE: VANM229.001CP1
CURRENT APPLICATION NUMBER: US/09/910,430
CURRENT FILING DATE: 2001-07-19

PRIOR APPLICATION NUMBER: PCT/BE00/00061
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: GB 9913425.6
PRIOR FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 27
LENGTH: 377
TYPE: PRT
ORGANISM: Ixodes ricinus
US-09-910-430-27

Query Match 44.9%; Score 44; DB 10; Length 377;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 ADLGSIASGRLYISRMK 20
DB 298 ADLGSIASGRLYISRMK 317

RESULT 10
US-10-157-031-355
Sequence 355, Application US/10157031
Publication No. US20030108890A1
GENERAL INFORMATION:
APPLICANT: Baranova, A. V.
APPLICANT: Yankovsky, N. K.
APPLICANT: Kozlov, A. P.
APPLICANT: Lobashov, A. V.
APPLICANT: Krukovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157,031
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 355
LENGTH: 391
TYPE: PRT
ORGANISM: Homo sapiens
US-10-157-031-355

Query Match 43.9%; Score 43; DB 9; Length 391;
Best Local Similarity 44.4%; Pred. No. 61;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 ADLGSIASGRLYISRM 18
DB 314 ADLGSIASGRLYISRM 331

RESULT 11
US-09-466-396A-110
Sequence 110, Application US/09466396A
Publication No. US20030119763A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.455C4
CURRENT APPLICATION NUMBER: US/09/466,396A
CURRENT FILING DATE: 1999-12-17
NUMBER OF SEQ ID NOS: 224
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 110
LENGTH: 391
TYPE: PRT
ORGANISM: Homo sapien
US-09-466-396A-110

Query Match 43.9%; Score 43; DB 9; Length 391;
Best Local Similarity 44.4%; Pred. No. 61;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
OY 1 ADLGSIASGRLYISRM 18
DB 314 ADLGSIASGRLYISRM 331

RESULT 12
US-09-735-705-110
Sequence 110, Application US/09735705
Patent No. US20020052329A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C14
CURRENT APPLICATION NUMBER: US/09/735,705
CURRENT FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 419
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 110
LENGTH: 391
TYPE: PRT
ORGANISM: Homo sapien
US-09-735-705-110

Query Match 43.9%; Score 43; DB 10; Length 391;
Best Local Similarity 44.4%; Pred. No. 61;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 ADLGSIASGRLYISRM 18
DB 314 ADLGSIASGRLYISRM 331

RESULT 13
US-09-850-716A-110
Sequence 110, Application US/09850716A
Patent No. US20020115139A1
GENERAL INFORMATION:
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Retter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C15
CURRENT APPLICATION NUMBER: US/09/850,716A
CURRENT FILING DATE: 2001-05-07
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 110
LENGTH: 391
TYPE: PRT
ORGANISM: Homo sapien
US-09-850-716A-110

Query Match 43.9%; Score 43; DB 10; Length 391;
Best Local Similarity 44.4%; Pred. No. 61;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
OY 1 ADLGSIASGRLYISRM 18

Db 314 ADYSGMSSGSLVYAKFL 331

RESULT 14
US-09-897-778-110
Sequence 110, Application US/09897778
Patent No. US20020147143A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Marnerakis, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Matanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Peckham, David W.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C16
CURRENT APPLICATION NUMBER: US/09/897,778
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 467
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 110
LENGTH: 391
TYPE: PRT
ORGANISM: Homo sapiens
US-09-897-778-110

Query Match 43.9%; Score 43; DB 10; Length 391;
Best Local Similarity 44.4%; Pred. No. 61;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 ADLSGASGRLYISRMK 18
|||::|||::|
314 ADYSGMSSGSLVYAKFL 331

Db

RESULT 15
US-09-902-684-2
Sequence 2, Application US/09902684
Patent No. US20020127640A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
INHIBITOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/902,684
FILING DATE: 12-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/026,408
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/024,056
FILING DATE: 16-Aug-1996
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0300002

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 392 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-902-684-2

Query Match 43.9%; Score 43; DB 10; Length 392;
Best Local Similarity 42.1%; Pred. No. 61;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 DLGSIASGRLYISRMK 20
|||||::|::|
Db 316 DLGSISSSEYVSVYQK 334

Search completed: July 11, 2003, 12:37:33
Job time: 5.86279 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 4.0326 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-6

Perfect score: 98

Sequence: 1 ADLSCIASGRLYISRMWGK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues.

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	55.1	191	2	S49162
2	48	49.0	591	2	A75474
3	46	46.9	524	2	B64302
4	45	45.9	406	2	A39339
5	45	45.9	957	2	T10633
6	44.5	45.4	434	2	C36144
7	44	44.9	205	2	D69804
8	44	44.9	222	2	C72657
9	44	44.9	453	2	AD1556
10	44	44.9	453	2	AF1198
11	44	44.9	870	2	T10634
12	43.5	44.4	1805	2	T02712
13	43	43.9	310	2	J00748
14	43	43.9	371	2	G84398
15	43	43.9	379	2	A42421
16	43	43.9	391	2	JC7118
17	43	43.9	423	2	B72528
18	43	43.9	622	2	A64494
19	43	43.9	627	2	G81719
20	43	43.9	628	2	G71565
21	42.5	43.4	203	2	D87567
22	42	42.9	126	2	G71269
23	42	42.9	311	2	D97450
24	42	42.9	311	2	AF2668
25	42	42.9	341	2	D98296
26	42	42.9	341	2	AD2887
27	42	42.9	379	2	S27383
28	42	42.9	396	2	I50707
29	42	42.9	417	2	S19724

30	42	42.9	700	2	D75508	serine/threonine p
31	42	42.9	740	2	T45091	catalase (EC 1.11.
32	42	42.9	1288	2	T09908	hypothetical prote
33	41	41.8	378	2	S38962	serpin - pig
34	41	41.8	383	2	S11433	ovalbumin - Japane
35	41	41.8	388	2	AF1907	hypothetical prote
36	41	41.8	397	2	I39184	bomapsin - human
37	41	41.8	418	2	JX0129	constrapsin precurs
38	41	41.8	433	1	ITRHC	alpha-1-antichymot
39	41	41.8	447	2	A72358	conserved hypotnet
40	41	41.8	485	2	C86336	hypothetical prote
41	41	41.8	487	2	T05679	H+-transporting tw
42	41	41.8	488	2	T43789	H+-transporting tw
43	41	41.8	492	2	A31886	H+-exporting ATPas
44	41	41.8	513	2	C44138	vacuolar H+-transp
45	41	41.8	513	2	A33281	H+-transporting tw

ALIGNMENTS

RESULT 1
S49162
ZG-21P protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 21-Jan-2000
C:Accession: S49162
R/Cronshagen, U.; Chen, C.; Kern, H.F.
A:Submitted to the EMBL Data Library, March 1994
A:Description: A novel protein expressed exclusively in pancreas is proposed to be a
A:Reference number: S49162
A:Accession: S49162
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-191 <CRO>
A:Cross-references: EMBL:Z30585; NID:9510193; PIDN:CAA83060.1; PID:9510194
C:Superfamily: antithrombin III

Query Match 55.1%; Score 54; DB 2; Length 191;
Best Local Similarity 57.9%; Pred. No. 0.3;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 2 DLSCIASGRLYISRMWGK 20
DB 102 DLSGITDSEELYSRAMQK 120

RESULT 2

A75474
serine proteinase, subtilase family - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
R/C:Accession: A75474

R/White, O.; Eisen, J.A.; Heldeberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 266, 1571-1577, 1999

A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MID:20036896; PMID:10567266

A:Accession: A75474
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-591 <WHI>
A:Cross-references: GB:AE001935; GB:AE000513; NID:96458517; PIDN:AAF10389.1; PID:9645
A:Experimental source: strain R1

C:Genetics:
A:Gene: DR0812
A:Map position: 1

C:Superfamily: subtilisin; subtilisin homology

Query Match 49.0%; Score 48; DB 2; Length 591;
Best Local Similarity 47.1%; Pred. No. 9.1;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 3 LSGIAGSGRLYSRMKG 19
 DB 255 VCGVAGSVNLYMARVLG 271

RESULT 3

B64302
 hypothetical protein M70018 - Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Oct-1999

C:Accession: B64302

R:Smith, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Nelson, C.F.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Kleen, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C.

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; M01D:96337999; PMID:8688087

A:Accession: B64302

A:Molecule type: DNA

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 1-524 <BUL>

A:Cross-references: GB:U67460; GB:L77117; NID:91550820; PIDN:AA09002.1; PID:91498776; T

C:Genetics:

A:Map position: FOR17658-19232

Query Match

Best local similarity 46.9%; Score 46; DB 2; Length 524;

Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 3 LSGIAGSGRLYSRMKG 19

DB 384 IGNIAGGKLYDNVIG 400

RESULT 4

A39339
 protein C inhibitor precursor - human

N:Alternate names: GP43; plasminogen activator inhibitor 3 precursor (PAI3); protein DKF

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Jun-2000

C:Accession: A39339; A37341; A26139; G02081; A60454; A60462; S41213; T12502; S41214; S41

J. Biol. Chem. 266, 15028-15034, 1991

A:Title: Organization of the gene coding for human protein C inhibitor (plasminogen acti

A:Reference number: A39339; M01D:91332018; PMID:1714450

A:Accession: A39339

A:Molecule type: DNA

A:Residues: 1-406 <MEI1>

A:Cross-references: GB:M68516; GB:M64881; NID:9189677; PIDN:AAA02811.1; PID:9189678; GB:

A:Molecule type: mRNA
 A:Residues: 1-54, 'V', '56-104', 'E', '106-406' <RAD>
 A:Cross-references: EMBL:U5464; NID:91144560; PIDN:AA060386.1; PID:91144561
 A:Experimental source: allelic variant B

R:Laurell, M.; Stenflo, J.

Thromb. Haemost. 62, 885-891, 1989

A:Title: Protein C inhibitor from human plasma: characterization of native and cleave

A:Reference number: A60434; M01D:90085156; PMID:2556811

A:Accession: A60434

A:Molecule type: protein

A:Residues: 20-39 <LAU>

R:Mizuta, K.; Yasuda, T.; Kishi, K.

Biochem. Genet. 27, 731-743, 1989

A:Title: Biochemical and genetic studies on GP43, a 43-KD glycoprotein detected immun

A:Reference number: A60462; M01D:90179688; PMID:2627198

A:Accession: A60462

A:Molecule type: protein

A:Residues: 377-386 <MI2>

A:Note: this protein was found to be polymorphic in a Japanese test population

R:Christenson, A.; Lilja, H.

Eur. J. Biochem. 220, 45-53, 1994

A:Title: Complex formation between protein C inhibitor and prostate-specific antigen

A:Reference number: S41212; M01D:94164172; PMID:7509746

A:Accession: S41213

A:Molecule type: protein

A:Residues: 29-38, 'X', '40-44', 'XX', '47-50', 'X', '52, 377', 'X', '379-398', 'X', '400' <CHR>

R:Pouskka, A.; Kleen, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z17525

A:Accession: T12502

A:Molecule type: mRNA

A:Residues: 'E', '29-406' <POU>

A:Cross-references: EMBL:AL080185; NID:95262662; PIDN:CA045766.1; PID:95262663

A:Experimental source: adult testis; clone DKFZP434P131

C:Comment: Inhibits protein C (see PIR:KXHU) and plasminogen activator (see PIR:UKHU

C:Genetics:

A:Gene: GDB:PCI

A:Cross-references: GDB:134739; OMIM:601841

A:Map position: 14q32.1-14q32.1

A:Introns: 207/1; 297/2; 346/3

A:Note: DKFZP434P131.1

C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; monomer; polymorphism; serine protease inhibitor
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-406/Product: protein C inhibitor #status experimental <MAT>
 F:249-262, 338/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:376/Inhibitory site: Arg (plasminogen activator) #status predicted

Query Match

Best local similarity 45.9%; Score 45; DB 2; Length 406;

Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 ADLSGIAGSGRLYSRMKG 20

DB 331 ADLSGIAGSGRLYSRMKG 350

RESULT 5

T10633
 hypothetical protein T13K14.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 18-Aug-2000

C:Accession: T10633

R:Bevan, M.; Pohl, T.; Weizengger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lem

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16991

A:Accession: T10633

A:Molecule type: DNA

A:Residues: 1-957 <BEV>

A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.70

A:Experimental source: cultivar Columbia; BAC clone T13K14

C:Genetics:

A:Gene: ATSP:T13K14.70

C:Genetics:
A:Gene: pstC1

Query Match 43.9%; Score 43; DB 2; Length 371;
Best Local Similarity 45.0%; Pred. No. 36;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMGR 20
| | | | | : | : | | |
Db 271 AGLSGVSAGVLLGVRAME 290

RESULT 15

A42421

Leukocyte elastase inhibitor - horse

N:Alternate names: plasminogen activator inhibitor-2 homolog

C:Species: Equus caballus (domestic horse)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 02-Jun-2000

C/Accession: A42421; A37276; S34062

R:Dublin, A.; Travis, J.; Enghild, J.J.; Potempa, J.

J. Biol. Chem. 267, 6576-6583, 1992

A>Title: Equine leukocyte elastase inhibitor. Primary structure and identification as a

A:Reference number: A42421; MUID:92202200; PMID:151869

A:Accession: A42421

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-379 <DUB1>

A:Cross-references: PIDN:AB21885.1; PID:9247842

A:Experimental source: leukocyte

A>Note: sequence extracted from NCBI backbone (NCBIP:89849)

R:Dublin, A.; Travis, J.; Enghild, J.J.; Potempa, J.

submitted to the Protein Sequence Database, December 1991

A:Reference number: A37276

A:Accession: A37276

A:Molecule type: protein

A:Residues: 1-41, 'E', '43-325, 'VD', 326-379 <DUB2>

R:Kordula, T.; Dublin, A.; Scholtink, H.; KoJ, A.; Heinrich, P.C.; Rose-John, S.

Biochem. J. 293, 187-193, 1993

A>Title: Molecular cloning and expression of an intracellular serpin: an elastase inhibi

A:Reference number: S34062; MUID:93319507; PMID:7687128

A:Accession: S34062

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-379 <KOR>

A:Cross-references: GB:M91161; NID:q164240; PIDN:AA97513.1; PID:q164241

C:Superfamily: antithrombin III

C:Keywords: serine proteinase inhibitor

Query Match 43.9%; Score 43; DB 2; Length 379;
Best Local Similarity 40.0%; Pred. No. 37;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMGR 20
| | | | | : | : | | |
Db 302 ADLSGMSGARDLFVSKTIHK 321

Search completed: July 11, 2003, 12:00:39
Job time : 5.13326 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 1.91268 Seconds

(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-6

Perfect score: 98

Sequence: 1 ADLSCGASGRLYSRMNGK 20

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

112892

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	91.8	380	1 SPB7_HUMAN	O75635 homo sapien
2	53	54.1	405	1 SPI2_MOUSE	O91x88 mus musculu
3	51	52.0	378	1 PT16_BOVIN	O02739 bos taurus
4	47	48.0	405	1 SB12_HUMAN	O96663 homo sapien
5	46	46.9	524	1 Y018_MERJA	O60334 methanococc
6	45	45.9	406	1 IPSE_HUMAN	P05134 homo sapien
7	44.5	45.4	433	1 COB2_PSEDE	P21632 pseudomonas
8	43	43.9	310	1 DP3B_MICIU	P21174 micrococcus
9	43	43.9	379	1 ILEU_HORSE	P05619 equus caball
10	43	43.9	391	1 SB13_HUMAN	O91x78 homo sapien
11	43	43.9	405	1 SPI2_HUMAN	O75830 homo sapien
12	43	43.9	622	1 YF54_MERJA	O38949 methanococc
13	42	42.9	126	1 RBFA_TREPA	O83860 treponema p
14	42	42.9	379	1 ILEU_HUMAN	P30740 homo sapien
15	42	42.9	396	1 SX11_CHICK	P48435 gallus gall
16	42	42.9	417	1 KBP_MOUSE	P29621 mus musculu
17	42	42.9	739	1 CAT4_MYCSM	O59557 mycobacteri
18	41	41.8	378	1 ILEU_PIG	P80229 sus scrofa
19	41	41.8	382	1 OVAL_COTJA	P19140 coturnix co
20	41	41.8	386	1 VAB2_GOSHI	O43473 gossypium h
21	41	41.8	392	1 SB11_HUMAN	O96615 homo sapien
22	41	41.8	397	1 SB10_HUMAN	P48595 homo sapien
23	41	41.8	413	1 HP55_TAMST	O09055 tamias sibi
24	41	41.8	418	1 COTR_MOUSE	P07759 mus musculu
25	41	41.8	423	1 AACT_HUMAN	P01011 homo sapien
26	41	41.8	433	1 VAB2_HORVU	O40079 hordeum vul
27	41	41.8	488	1 VAB1_GOSHI	O43432 gossypium h
28	41	41.8	488	1 VAB1_HORVU	O40078 hordeum vul
29	41	41.8	492	1 VAB2_AARTH	P11574 arabidopsis
30	41	41.8	513	1 VAB1_BOVIN	P31407 bos taurus
31	41	41.8	513	1 VAB1_HUMAN	P15313 homo sapien
32	41	41.8	1322	1 SAL1_MOUSE	O98f14 mus musculu
33	40	40.8	283	1 PSTA_RHILU	O98f14 rhizobium l

34	40	40.8	328	1 GRHR_HORSE	O18821 equus caball
35	40	40.8	328	1 GRHR_PIG	P49922 sus scrofa
36	40	40.8	378	1 PT16_MOUSE	O60854 mus musculu
37	40	40.8	420	1 DCDA_AOUAE	O67262 aquifex aeo
38	40	40.8	453	1 VAB2_CHICK	P49712 gallus gall
39	40	40.8	485	1 VAB2_TRYCO	O26976 trypanosoma
40	40	40.8	490	1 VAB2_DROME	P31409 drosophila
41	40	40.8	491	1 VAB2_CAEEL	O19626 caenorhabdit
42	40	40.8	492	1 OPU2_BACSU	O06493 bacillus su
43	40	40.8	492	1 VAB1_ACETAT	O38681 acetabulari
44	40	40.8	492	1 VAB2_ACETAT	O38680 acetabulari
45	40	40.8	494	1 VAB2_HELVI	P31410 heliothis v

ALIGNMENTS

RESULT 1
SPB7_HUMAN STANDARD: PRT: 380 AA.
AC O75635;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Megsin (TP55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97326116; PubMed=9182567;
RA Tsujimoto M., Tsunooka N., Ishida N., Kurihara T., Iwasa F.,
Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
Katayama T., Nakao M., Yamachi K., Hashino J., Hanyama M., Miura K.,
Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.;
RT "Purification, cDNA cloning, and characterization of a new serpin with
megakaryocyte maturation activity."
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98376492; PubMed=9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, megin, is a new serpin upregulated in
IgA nephropathy."
RL J. Clin. Invest. 102:828-836(1998).
CC -!- FUNCTION: Might function as an inhibitor of lys-specific
proteases. Might influence the maturation of megakaryocytes via
its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: D88575; BAA31232.1; -
DR EMBL: AF027866; AAC64506.1; -
DR HSSP: P05619; IHLF.
DR Genew: HGNC:13902; SERPINB7.
DR KIM: 603357; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN.1.

KW Serpin; Serine protease inhibitor.
 FT ACT SITE 347 348 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 4290 MW; 9A2CDB6C3CFF605 CRC64;

Query Match 91.8%; Score 90; DB 1; Length 380;
 Best Local Similarity 95.0%; Pred. No. 2.5e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMGK 20
 DB 305 ADLSGIASGRLYISRMGK 324

RESULT 2
 ID SPI2_MOUSE STANDARD; PRT; 405 AA.
 AC O9JX88; Q9D955; Q9D8Z3;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serpin I2 precursor (Serine protease inhibitor 14).
 GN SERPINI2 OR SPI14.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.
 RC TISSUE-Pancreas;
 RA Chang W.S., Lin S.C., Wu C.W.;
 RT "Isolation and characterization of mouse pancreas-specific serpin
 gene.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pancreas;
 RX MEDLINE=2108560; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakura I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
 RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hornum M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Noridone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wymshay-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: Expressed in pancreas.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL AF251276; AAF65821.1; -;
 DR EMBL AK007347; BAB24976.1; -;
 DR EMBL AK007510; BAB25079.1; -;
 DR HSSP; P01008; IATH.

DR MGD; MG1:894672; SPI14.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin. 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Glycoprotein; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 405
 FT ACT SITE 357 358 REACTIVE BOND (BY SIMILARITY).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 5 5 I -> M (IN REF. 2; BAB25079).
 FT CONFLICT 195 195 K -> T (IN REF. 2; BAB24976).
 FT CONFLICT 207 207 D -> S (IN REF. 2; BAB24976).
 FT CONFLICT 212 218 KYPMKKA -> RVPPTDEV (IN REF. 2; BAB24976).
 FT CONFLICT 246 246 F -> Y (IN REF. 2; BAB24976).
 SQ SEQUENCE 405 AA; 45775 MW; 04DF38BE8545DEA8 CRC64;

Query Match 54.1%; Score 53; DB 1; Length 405;
 Best Local Similarity 52.6%; Pred. No. 0.31;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 DLSGIASGRLYISRMGK 20
 DB 316 DLSGITDSEVYRVYMQK 334

RESULT 3
 ID PTI6_BOVIN STANDARD; PRT; 378 AA.
 AC 002739;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine proteinase inhibitor B-43.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97165893; PubMed=9013786;
 RA Nakaya N., Nishibori M., Kawabata M., Saeki K.;
 RT "Cloning of a serine proteinase inhibitor from bovine brain:
 RT expression in the brain and characterization of its target
 RT proteinases.";
 RL Brain Res. Mol. Brain Res. 42:293-300(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES
 CC PRESENT IN THE BRAIN OR EXTRA-CELLULAR FLUID.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL D55670; BAA19875.1; -;
 DR HSSP; P05120; IBA7.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin. 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT SITE 343 344 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 378 AA; 42560 MW; 66AF49CCCE263A CRC64;

Query Match 52.0%; Score 51; DB 1; Length 378;

Best Local Similarity 50.0%; Pred. No. 0.62;
Matches 10; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 ADLSGASGRGLYISRMKG 20
DB 301 ADFSGMSSGRGLHLSKVMHK 320

RESULT 4

SB12_HUMAN STANDARD; PRT; 405 AA.
AC 096P63;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Serpin B12.
GN SERPINB12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA MEDLINE-21623618; PubMed-11604408;
RA Askev Y.S., Pak S.C., Luke C.J., Askev D.J., Cataltepe S., Mills D.R.,
RA Kato H., Lehoccky J., Dewar K., Birren B., Silverman G.A.;
RT SERPINB12 is a novel member of the human ov-serpin family that is
RT widely expressed and inhibits trypsin-like serine proteinases.";
RL J. Biol. Chem. 276:49320-49330(2001).
CC 1- FUNCTION: Inhibits trypsin and plasmin, but not chymotrypsin,
CC coagulation factor Xa, or urokinase-type plasminogen activator.
CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC 1- TISSUE SPECIFICITY: Expressed in many tissues, including brain,
CC bone marrow, lymph node, heart, lung, liver, pancreas, testis,
CC ovary, and intestine.
CC 1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF411191; AAL05571.1;
DR Gene: HGNC:14220; SERPINB12.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR PROSITE: PS00284; SERPIN.1.
KW Serpin, Serine protease inhibitor.
FT ACT_SITE 370 371 REACTIVE_BOND (BY SIMILARITY).
SQ SEQUENCE 405 AA; 46276 MW; FFE12D4C9B7F3DFA CRC64;

Query Match 48.0%; Score 47; DB 1; Length 405;
Best Local Similarity 45.0%; Pred. No. 3;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGASGRGLYISRMKG 20
DB 328 ADLTGISPNLYLSKVIHK 347

RESULT 5

Y018_METJA STANDARD; PRT; 524 AA.
AC 060324;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M70018.
GN M70018.
OS Methanococcus jannaschii.

OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed-8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurd M.A., Kaine B.P., Bordovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67460; AAB98002.1;
DR TRIGR; M70018;
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 13 33 POTENTIAL.
SQ SEQUENCE 524 AA; 56565 MW; 0F878499773FF87C CRC64;

Query Match 46.9%; Score 46; DB 1; Length 524;
Best Local Similarity 41.2%; Pred. No. 5.7;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 3 LSGIAGSGRLYISRMKG 19
DB 384 IGNIAGGKLYDNYVG 400

RESULT 6

IPSP_HUMAN STANDARD; PRT; 406 AA.
ID IPSP_HUMAN
AC P05154; Q9UG30; Q07616;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma serine protease inhibitor precursor (PCI) (Protein C inhibitor)
DE (Plasminogen activator inhibitor-3) (PAI3) (Acrosomal serine protease
DE inhibitor).
GN SERPINAS OR PCI OR PLANH3 OR PROCI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-87109153; PubMed-3027058;
RA Suzuki K., Devashiki Y., Nishioke J., Kurachi K., Akira M.,
RA Yamamoto S., Hashimoto S.;
RT "Characterization of a cDNA for human protein C inhibitor. A new
RT member of the plasma serine protease inhibitor superfamily.";
RL J. Biol. Chem. 262:611-616(1987).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-91048502; PubMed=2173165;
RA Meijers J.C.M., Chung D.W.;
RT "Evidence for a glycine residue at position 316 in human protein C
RT inhibitor.";
RL Thromb. Res. 59:389-393(1990).
RN 13

RP SEQUENCE FROM N.A.
 RX MEDLINE-91332018; PubMed-1714450;
 RA Weijers J.C.M., Chung D.W.;
 RT "Organization of the gene coding for human protein C inhibitor
 (plasminogen activator inhibitor-3). Assignment of the gene to
 chromosome 14.";
 RL J. Biol. Chem. 266:15028-15034(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94198434; PubMed-8148499;
 RA Hayashi T., Suzuki K.;
 RT "Gene organization of human protein C inhibitor, a member of SERPIN
 family proteins encoded in five exons.";
 RL Int. J. Hematol. 58:213-224(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Liver;
 RA Radtke K.P., Greengard J.S., Fernandez J.A., Villoutreix B.O.,
 RL Griffin J.H.;
 RT Submitted (Jan-1996) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANTS G-44; V-55; S-64; V-94; E-105; P-115
 AND R-217.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RL Nicolson D.A.;
 RT Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Skin;
 RA Strausberg R.;
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 27-406 FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE-93329000; PubMed-8471250;
 RA Moore A., Penfold L.M., Johnson J.L., Latchman D.S., Moore H.D.;
 RT "Human sperm-egg binding is inhibited by peptides corresponding to
 core region of an acrosomal serine protease inhibitor.";
 RL Mol. Reprod. Dev. 34:280-291(1993).
 RN [9]
 RP SEQUENCE OF 28-406 FROM N.A.
 RC TISSUE-Testis;
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wilmann S.;
 RL Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE OF 20-39.
 RX MEDLINE-90085156; PubMed-2556811;
 RA Laurell M., Stenflo J.;
 RT "Protein C inhibitor from human plasma: characterization of native
 and cleaved inhibitor and demonstration of inhibitor complexes with
 plasma kallikrein.";
 RL Thromb. Haemost. 62:885-891(1989).
 RN [11]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-91046026; PubMed-2172989;
 RA Kuhn L.A., Griffin J.H., Fisher C.L., Greengard J.S., Bouma B.N.,
 RA Espana F., Tainer J.A.;
 RT "Elucidating the structural chemistry of glycosaminoglycan
 recognition by protein C inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8506-8510(1990).
 CC -1- FUNCTION: INHIBITS ACTIVATED PROTEIN C AS WELL AS PLASMINOGEN
 ACTIVATORS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER. SECRETED IN PLASMA.
 CC -1- DISEASE: A DEFICIENCY OF PROTEIN C INHIBITOR IS THE CAUSE OF A
 COMBINED FACTOR V/VIII DEFICIENCY DISEASE.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL, outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M68516; AAA02811.1; -
 DR EMBL: J02639; AAA35688.1; -
 DR EMBL: S69366; AAB30461.1; -
 DR EMBL: S69364; AAB30461.1; JOINED.
 DR EMBL: S69364; AAB30461.1; JOINED.
 DR EMBL: S69365; AAB30461.1; JOINED.
 DR EMBL: AF361796; AAK27240.1; -
 DR EMBL: U35464; AAB60386.1; -
 DR EMBL: BC008915; AAH08915.1; -
 DR EMBL: S58545; AAB26244.2; -
 DR EMBL: AL080185; CAB45766.1; -
 DR PIR: A39339; A39339.
 DR PIR: A39339; A39339.
 DR PDB: 1PAI; 15-OCT-94.
 DR PDB: 2PAI; 15-OCT-94.
 DR Genew; HGNC:8723; SERPIN5.
 DR MIM: 601841; -
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 3.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Plasma; Glycoprotein; Signal;
 KW Polymorphism; 3d-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 406
 FT CARBOHYD 249 249
 FT CARBOHYD 262 262
 FT CARBOHYD 338 338
 FT ACT_SITE 373 374
 FT VARIANT 44 44
 FT FT
 FT FT
 FT VARIANT 55 55
 FT FT
 FT FT
 FT VARIANT 64 64
 FT FT
 FT VARIANT 94 94
 FT FT
 FT VARIANT 105 105
 FT FT
 FT FT
 FT VARIANT 115 115
 FT FT
 FT VARIANT 217 217
 FT FT
 FT CONFLICT 28 28
 FT CONFLICT 221 221
 FT CONFLICT 335 335
 FT CONFLICT 384 384
 SQ SEQUENCE 406 AA; 45701 MW; 0B9D1A519341B8C9 CRC64;
 Query Match 45.9%; Score 45; DB 1; Length 406;
 Best Local Similarity 45.0%; Pred. No. 6.4;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 ADLGIASGGRGLYSRMGK 20
 Db 331 ADLGIASGGRGLYSRMGK 350
 |||||: : : : :
 RESULT 7
 COBB_PSEDE STANDARD; PRT; 433 AA.
 ID COBB_PSEDE P21632;
 AC P21632;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cobyrynic acid A.C-diamide synthase.
 GN COBB.
 OS Pseudomonas denitrificans.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 NCBI_TaxID=43306;

of the reactive center." J. Biol. Chem. 263:7364-7369(1988).
 [4]
 X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 MEDLINE-92389327; PubMed-1518052;
 Baumann U., Bode W., Huber R., Travis J., Potempa J.;
 "Crystal structure of cleaved equine leucocyte elastase inhibitor
 determined at 1.95-A resolution.";
 J. Mol. Biol. 226:1207-1218(1992).
 -1- FUNCTION: THIS INHIBITOR IS THOUGHT TO BE INVOLVED IN THE CONTROL
 OF INTRACELLULAR PROTEIN TURNOVER. BINDS TO THYMOSIN BETA-4.
 CC OF INTRACELLULAR PROTEIN TURNOVER. BINDS TO THYMOSIN BETA-4.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M91161; AAA97513.1; -
 DR PIR; A28060; A28060.
 DR PIR; A37276; A37276.
 DR PIR; A42421; A42421.
 DR PIR; S34062; S34062.
 DR PDB; 1HE; 31-JAN-94.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 DR Serpin; Serine protease inhibitor; 3D-structure.
 KM MOD_RES 1 1
 FT ACT_SITE 344 345
 FT STRAND 2 22
 FT HELIX 28 30
 FT HELIX 32 44
 FT TURN 45 45
 FT HELIX 48 58
 FT TURN 59 59
 FT HELIX 60 62
 FT TURN 64 65
 FT HELIX 66 67
 FT TURN 78 78
 FT STRAND 85 95
 FT TURN 96 97
 FT HELIX 102 112
 FT STRAND 115 119
 FT TURN 121 123
 FT HELIX 125 139
 FT TURN 140 142
 FT TURN 150 151
 FT TURN 155 156
 FT STRAND 157 171
 FT HELIX 177 179
 FT STRAND 181 186
 FT STRAND 192 209
 FT HELIX 210 212
 FT TURN 213 213
 FT STRAND 214 221
 FT TURN 222 223
 FT STRAND 226 233
 FT HELIX 244 248
 FT TURN 249 249
 FT HELIX 252 259
 FT STRAND 261 263
 FT STRAND 265 274
 FT STRAND 276 283
 FT HELIX 285 291
 FT TURN 292 292
 FT HELIX 295 297
 FT TURN 299 301

FT HELIX 305 308
 FT STRAND 313 326
 FT STRAND 330 343
 FT STRAND 350 353
 FT STRAND 358 364
 FT TURN 365 368
 FT STRAND 369 376
 SQ SEQUENCE 379 AA; 43006 MW; 72DED1999371427A CRC64;
 Query Match 43 9%; Score 43; DB 1; Length 379;
 Best Local Similarity 40.0%; Pred. No. 13;
 Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 QY 1 ADLGIASGRLYISRMNGK 20
 DB 302 ADLSGMSGARDLFVSKIIHK 321
 RESULT 10
 ID SB13_HUMAN STANDARD; PRT; 391 AA.
 AC Q9UYV8; Q9UBW1; Q9UKG0; Q9HCX1;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Hurpin (Hacat UV-repressible serpin) (Protease inhibitor 13)
 DE (Headpin) (Serpin B13).
 GN SERPINB13 OR PIL3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RP [1]
 RC SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Skin;
 RX MEDLINE-99443898; PubMed-10512713;
 RA Abts H.F., Weiss T., Mirzomahmadsadeh A., Koehrer K., Michel G.,
 RT "Cloning and characterization of hurpin (Protease inhibitor 13): a new
 RT skin specific, UV-repressible serine proteinase inhibitor of the
 RT ovalbumin serpin family";
 RL J. Mol. Biol. 293:29-39(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE-99458661; PubMed-10527881;
 RA Spring P., Nakashima T., Frederick M., Henderson Y., Clayman G.;
 RT "Identification and cDNA cloning of headpin, a novel differentially
 RT expressed serpin that maps to chromosome 18q";
 RL Biochem. Biophys. Res. Commun. 264:299-304(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE-20461850; PubMed-11004515;
 RA Nakashima T., Pak S.C., Silverman G.A., Spring P.M., Frederick M.J.,
 RA Clayman G.L.;
 RT "Genomic cloning, mapping, structure and promoter analysis of HEADPIN,
 RT a serpin which is down-regulated in head and neck cancer cells";
 RL Biochim. Biophys. Acta 1492:441-446(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RP TISSUE=Placenta;
 RX MEDLINE-21213570; PubMed-11313015;
 RA Abts H.F., Weiss T., Schreuring S., Scott F.L., Irving J.A., Michel G.,
 RA Bird P.I., Ruzicka T.;
 RT "Sequence, organization, chromosomal localization and alternative
 RT splicing of the human serine protease inhibitor gene hurpin (PIL3),
 RT which is up-regulated in psoriasis";
 RL DNA Cell Biol. 20:123-131(2001).
 CC -1- FUNCTION: May play a role in the proliferation or differentiation
 CC of keratinocytes.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SKIN-SPECIFIC.

CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: AJ001696; CA004935.2; -
DR EMBL: AJ001697; CA004936.2; -
DR EMBL: AJ001698; CA004937.1; -
DR EMBL: AF169949; AAD55765.1; -
DR EMBL: AF216854; AAF72879.1; -
DR EMBL: AJ278717; CAC03569.1; -
DR HSSP: P05120; 1B7.
DR Genew: HGNC:8944; SERPINB13.
DR MIM: 604445; -
DR InterPro: IPR000240; Maspin.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin. 1.
DR PRINTS: PRO0676; MASPIN.
DR SMART: SM00093; SERPIN. 1.
DR PROSITE: PS00284; SERPIN. 1.
KW Serpin; Serine protease inhibitor; Alternative splicing.
FT ACT_SITE 356 357 REACTIVE_BOND (BY SIMILARITY).
FT VARSPIC 206 257 MISSING (IN ISOCORE 2).
FT CONFLICT 8 8 S -> N (IN REF. 4).
FT CONFLICT 75 75 MISSING (IN REF. 1; CA004937).
FT CONFLICT 293 293 G -> S (IN REF. 2).
FT CONFLICT 297 297 E -> Q (IN REF. 1; CA004937).
SQ SEQUENCE 391 AA: 44276 MW: 2CAB8558DBAC2B09 CRC64;
Query Match 43.9%; Score 43; DB 1; Length 391;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
DY 1 ADLSGIASGRLYISRM 18
Db 314 ADYSGMSGGLYAKFL 331
RESULT 11
SP12_HUMAN STANDARD; PRT; 405 AA.
AC 075830;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serpin I2 precursor (Myoepithelium-derived serine protease inhibitor)
DE (Pancrein) (Protease inhibitor 14) (TSA2004).
GN SERPIN2 OR P114 OR MEPI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed-9624529;
RA Okaiki K., Nagata M., Suzuki M., Fujiwara T., Miyoshi Y., Ishikawa O.,
RA Ohigashi H., Imoka S., Takahashi E., Nakamura Y.,
RT "Isolation and characterization of a novel human pancreas-specific
RL gene, pancpn, that is down-regulated in pancreatic cancer cells.";
RN Genes Chromosomes Cancer 22:179-185(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed-10097100;
RA Xiao G., Liu Y.E., Gentz R., Sang Q.A., Ni J., Goldberg I.D.,
RA Shi Y.E.;
RT "Suppression of breast cancer growth and metastasis by a serpin
RT myoepithelium-derived serine proteinase inhibitor expressed in the
RT mammary myoepithelial cells.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:3700-3705(1999).
CC -1- TISSUE SPECIFICITY: Expressed in pancreas and adipose tissues.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL: AB006423; BAA33766.1; -
DR EMBL: AF130470; AAD34723.1; -
DR HSSP: P05120; 1B7.
DR Genew: HGNC:8945; SERPINI2.
DR MIM: 605587; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin. 1.
DR SMART: SM00093; SERPIN. 1.
DR PROSITE: PS00284; SERPIN. 1.
KW Serpin; Serine protease inhibitor; Glycoprotein; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 405 SERPIN I2.
FT ACT_SITE 357 358 REACTIVE_BOND (BY SIMILARITY).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 405 AA: 46145 MW: 5BA18C60E4FDE9A4 CRC64;
Query Match 43.9%; Score 43; DB 1; Length 405;
Best Local Similarity 42.1%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
DY 2 DLGSIASGRLYISRMGX 20
Db 316 DLGSISSSEVYVSQYTK 334
RESULT 12
YF54_METJA STANDARD; PRT; 622 AA.
AC 058949;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ1554 precursor.
GN MJ1554.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-9637999; PubMed-8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uitterlind T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Bordovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO A.FULGIDUS AF0817.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U67596; AAB99574.1; -

DR TIGR: M01554; -

KW Hypothetical protein; Signal; Complete proteome.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 622 HYPOTHETICAL PROTEIN M1554.

SO SEQUENCE 622 AA; 72282 MW; 86CA6A62F3854D73F CRC64;

QY Query Match 43.9%; Score 43; DB 1; Length 622;

Best Local Similarity 47.1%; Pred. No. 21;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

3 LSGIASGRLYSRMGK 19

DB 428 LTGLKGERIYAVRFMG 444

RESULT 13

RBFA_TREPA STANDARD; PRT; 126 AA.

AC 08360; 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ribosome-binding factor A.

CN RBFA OR TP0890.

OS Treponema pallidum.

OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_TaxID=160;

RA [1]

RC SEQUENCE FROM N.A.

RE MEDLINE=96832770; PubMed=9665876;

RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,

RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,

RA Khalak H., Richardson D., Howell J.K., Childsbarrow M., Uitterback T.,

RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,

RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,

RA Venter J.C.;

RT "Complete genome sequence of Treponema pallidum, the syphilis

spirochete.";

RL Science 281:375-388(1998).

CC -1- FUNCTION: ASSOCIATES WITH FREE 30S RIBOSOMAL SUBUNITS (BUT NOT

WITH 30S SUBUNITS THAT ARE PART OF 70S RIBOSOMES OR POLYSOMES).

CC ESSENTIAL FOR EFFICIENT PROCESSING OF 16S RNA. MAY INTERACT WITH

THE 5'TERMINAL HELIX REGION OF 16S RNA (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).

CC -1- SIMILARITY: BELONGS TO THE RBFA FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE001258; AAC26580.1; -

DR TIGR: TP0890; -

DR InterPro: IPR000238; Rib_bind_facta.

DR Pfam: PF02033; RBFA.1.

DR ProDom: PD007327; Rib_bind_facta; 1.

DR TIGRFAMs: TIGR00082; rbfA; 1.

DR PROSITE: PS01319; RBFA.1.

DR RNA processing: Complete proteome.

SO SEQUENCE 126 AA; 14127 MW; DAB19B9861BBD2B3 CRC64;

Query Match 42.9%; Score 42; DB 1; Length 126;

Best Local Similarity 47.4%; Pred. No. 6;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 LLSGIASGRLYSRMGK 20

DB 44 DVSGGVCAHYVSSFMGK 62

RESULT 14

ILEU_HUMAN STANDARD; PRT; 379 AA.

AC P30740; 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase

inhibitor) (M/NEI) (EI).

CN SERPINB1 OR ELANH2 OR PI2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RA [1]

RC SEQUENCE FROM N.A.

RE MEDLINE=92302296; PubMed=1376927;

RA Remold-O'Donnell E., Chin J., Alberts M.;

RA "Sequence and molecular characterization of human monocyte/neutrophil

elastase inhibitor";

RL Proc. Natl. Acad. Sci. U.S.A. 89:5635-5639(1992).

CN [2]

OS SEQUENCE FROM N.A.

RE MEDLINE=9829265; PubMed=9630619;

RA Zeng W., Silverman G.A., Remold-O'Donnell E.;

RA "Structure and sequence of human M/NEI (monocyte/neutrophil elastase

inhibitor), an OY-serpin family gene.";

RL Gene 213:179-187(1998).

CN [3]

RC SEQUENCE FROM N.A.

RE TISSUE=Cervix;

RA Strausberg R.;

RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

RA [4]

RE SEQUENCE OF 178-185; 204-210 AND 364-371.

RA TISSUE=Keratinocytes;

RE MEDLINE=93162043; PubMed=1286667;

RA Rasmussen H.H., Van Damme J., Puype M., Gesser B., Celis J.E.,

RA Vandekekerckhove J.;

RT "Microsequences of 145 proteins recorded in the two-dimensional gel

protein database of normal human epidermal keratinocytes.";

RL Electrophoresis 13:960-969(1992).

CC -1- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASES

ELASTASE, CATHEPSIN G AND PROELINASE-3.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OY-SERPIN SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M93056; -; NOT ANNOTATED_CDS.

DR EMBL: AF053630; AAC31394.1; -

DR EMBL: BC009015; AAH09015.1; -

DR PIR: S27383; S27383.

DR HSP: P05619; ILE.

DR Aarhus/Sheat-2DPAGE: 4314; IEF.

DR Genew: HGNC:3111; SERPINB1.

DR MIM: 130135; -

DR InterPro: IPR000215; Serpin.

DR Pfam: PF00079; serpin; 1.

DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KM Serpin; Serine protease inhibitor.
 FT ACT_SITE 344 345
 SQ SEQUENCE 379 AA; 42741 MW; BA608DFC8D8C3 CRC64;

Query Match

Best Local Similarity 42.9%; Score 42; DB 1; Length 379;

Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 ADLSGIASGRLYISRMNGK 20

DB 302 ADLSGMSGARDIFISKIVHK 321

RESULT 15

SK11_CHICK STANDARD; PRT; 396 AA.
 ID SK11_CHICK
 AC P48435;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor SOX-11.
 GN SOX11.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=95267693; PubMed=7748786;
 RA Uvanogho D., Rex M., Cartwright E.J., Pearl G., Healy C.,
 RT Scotting P.J., Sharpe P.T.;
 RT "Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes
 suggests an interactive role in neuronal development.";
 RL Mech. Dev. 49:23-36(1995).
 CC -1- FUNCTION: MAY FUNCTION AS SWITCHES IN NEURONAL DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: LOW LEVEL EXPRESSION IS SEEN IN
 CC UNDIFFERENTIATED PROLIFERATING CELLS OF NEURAL EPITHELIUM. A
 CC GREATER EXPRESSION IS SEEN IN THE MATURING NEURONS AFTER THEY
 CC LEAVE THE NEURAL EPITHELIUM. IT IS ALSO FOUND IN THE GUT
 CC EPITHELIUM AND ADRENAL MEDULLA.
 CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U12534; AAB09664.1; -
 DR HSSP: P48436; 1SX9
 DR InterPro: IPR000910; HMG_12_box.
 DR Pfam: PF00505; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 KM Transcription regulation; DNA-binding; Nuclear protein.
 FT DNA_BIND 49 117
 FT DOMAIN 162 165 HMG_BOX
 FT DOMAIN 204 212 POLY-ALA.
 FT DOMAIN 288 294 POLY-GLU.
 FT DOMAIN 332 335 POLY-PRO.
 FT DOMAIN 332 335 POLY-ARG.
 SQ SEQUENCE 396 AA; 43503 MW; 8E4B0A457F8BA83 CRC64;

Query Match

Best Local Similarity 42.9%; Score 42; DB 1; Length 396;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLSGIASGRLYIS 15

DB 266 EVRGAGGRLYYS 279

Search completed: July 11, 2003, 11:55:17
 Job time : 2.91268 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 8.73181 Seconds
(without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-6
Perfect score: 98
Sequence: 1 ADLGGIAGSGRLYSRMWCK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORNITHINE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	84.7	360	11	Q9D695
2	83	84.7	360	11	Q9D695
3	54	55.1	191	11	Q63547
4	50	51.0	615	4	Q9NXX3
5	50	51.0	615	4	Q8TAP0
6	48	49.0	423	11	Q9D7P9
7	48	49.0	531	16	Q9RW57
8	46	46.9	377	6	Q9M2I6
9	46	46.9	377	11	Q08804
10	46	46.9	400	5	Q9GPI2
11	46	46.9	400	5	Q9GPI2
12	45	45.9	373	5	Q9N865
13	45	45.9	385	5	Q26509
14	45	45.9	942	10	Q65309
15	45	45.9	942	10	Q9C508
16	45	45.9	942	10	Q945R3

17	45	45.9	957	10	Q9SUC2	Q9SUC2 arabidopsis
18	44	44.9	205	16	Q31560	Q31560 bacillus su
19	44	44.9	222	17	Q9YE87	Q9YE87 aeropyrum p
20	44	44.9	377	5	Q9GPI3	Q9GPI3 ixodes ricl
21	44	44.9	379	11	Q9D7S8	Q9D7S8 mus musculu
22	44	44.9	379	11	Q9D154	Q9D154 mus musculu
23	44	44.9	386	13	Q73860	Q73860 meleagris g
24	44	44.9	388	11	Q9COV3	Q9COV3 mus musculu
25	44	44.9	388	11	Q91212	Q91212 mus musculu
26	44	44.9	453	16	Q9ZD31	Q9ZD31 listeria in
27	44	44.9	453	16	Q8Y8B8	Q8Y8B8 listeria in
28	44	44.9	870	10	Q9SUC1	Q9SUC1 arabidopsis
29	43.5	44.4	293	2	Q9LBR0	Q9LBR0 porphyromon
30	43.5	44.4	402	12	Q910Z6	Q910Z6 rana tigrin
31	43.5	44.4	431	12	Q9YKX2	Q9YKX2 epizootic h
32	43.5	44.4	431	12	Q8Q202	Q8Q202 rana tigrin
33	43.5	44.4	1805	10	Q81060	Q81060 arabidopsis
34	43	43.9	203	6	Q9GMA8	Q9GMA8 sus scrofa
35	43	43.9	231	6	Q28500	Q28500 macaca mula
36	43	43.9	309	17	Q97B69	Q97B69 thermoplasma
37	43	43.9	369	17	Q8TJ67	Q8TJ67 methanocarc
38	43	43.9	371	17	Q9HML6	Q9HML6 halobacteri
39	43	43.9	378	10	Q8S2S5	Q8S2S5 thelunglei
40	43	43.9	382	11	Q8VHP7	Q8VHP7 mus musculu
41	43	43.9	416	16	Q98KS9	Q98KS9 rhizobium l
42	43	43.9	418	11	Q62257	Q62257 mus musculu
43	43	43.9	423	17	Q919U0	Q919U0 aeropyrum p
44	43	43.9	477	5	Q8SR34	Q8SR34 encephalito
45	43	43.9	484	10	Q9LU12	Q9LU12 blastocysti

ALIGNMENTS

RESULT 1	ID	Q9D695	PRELIMINARY;	PRT;	360 AA.
AC	Q9D695	01-JUN-2001 (TREMBL)	17, Created		
DT	01-JUN-2001 (TREMBL)	17, Last sequence update)			
DT	01-JUN-2002 (TREMBL)	21, Last annotation update)			
DE	4531416M05R1K	protein (Megs1)			
GN	SERPINB7 OR 4531416M05RIK				
OS	Mus musculus (mouse)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SRRATN-C57B/6J; TISSUE=SKIN;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochua H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carinici P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gastlisch S., Hill D., Hofmann M., Hume D.A., Kamlyk M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,				
RA	Hayashizaki Y.,				
RT	"Functional annotation of a full-length mouse cDNA collection.";				
RL	Nature 409:685-690(2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				

```

RX MEDLINE-21368006; PubMed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent meginin revealed its up-regulation in
RT mesangiolipoliferative nephritis."
RL Kidney Int. 60:641-652(2001).
DR EMBL: AK014524; BAB29410.1; -.
DR EMBL: AF105328; AAL16768.1; -.
DR HSSP: P05121; 1A7C.
DR MGD: MGI:2151053; Serpinb7.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCF93CF4 CRC64;

Query Match
Best Local Similarity 84.7%; Score 83; DB 11; Length 380;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLGSIGSGRLYSRMKG 20
DB 305 ADLGSIGSGRLYSKLMHK 324

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
AC Q920J5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Meginin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21368006; PubMed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent meginin revealed its up-regulation in
RT mesangiolipoliferative nephritis."
RL Kidney Int. 60:641-652(2001).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AF105329; AAL16769.1; -.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
KW Serpin.
SQ SEQUENCE 380 AA; 42821 MW; D8076CA8EE2C2FBC CRC64;

Query Match
Best Local Similarity 84.7%; Score 83; DB 11; Length 380;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLGSIGSGRLYSRMKG 20
DB 305 ADLGSIGSGRLYSKLMHK 324

RESULT 3
Q63547 PRELIMINARY; PRT; 191 AA.
AC Q63547;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 2G-21P.
GN 2G-21.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-PANCREAS;
RA Cronshagen U., Chen C., Kern H.F.;
RT "A novel protein expressed exclusively in pancreas is proposed to be a
RT serpin."
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: Z30585; CAA83060.1; -.
DR HSSP: P01008; 1A7H.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
KW Serpin.
SQ SEQUENCE 191 AA; 21825 MW; 9EC312FEDBD169320 CRC64;

Query Match
Best Local Similarity 55.1%; Score 54; DB 11; Length 191;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 DLGSIASGRLYSRMKG 20
DB 102 DLGSIIDSSSELYSRAMOK 120

RESULT 4
Q9N0X3 PRELIMINARY; PRT; 615 AA.
AC Q9N0X3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CDNA FLJ11078 f1s, clone PLACE1005102, weakly similar to RING CANAL
DE protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;
RT "NEO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK001940; BAA91990.1; -.
DR HSSP: Q05516; 1CS3.
DR InterPro: IPR000210; BTB_POZ.
DR Pfam: PF00651; BTB; 1.
DR Pfam: PF01344; Kelch; 5.
DR SMART: SM00225; BTB; 1.
DR PROSITE: PS0097; BTB; 1.
SQ SEQUENCE 615 AA; 68139 MW; 1583474AF5A45EBD CRC64;

Query Match
Best Local Similarity 51.0%; Score 50; DB 4; Length 615;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 SGIASGRLYS 15
DB 454 AGAASGRLYS 465

RESULT 5
Q8TAP0 PRELIMINARY; PRT; 615 AA.
AC Q8TAP0;

```

```

DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 68.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC026319; AAR26319.1; -.
KW Hypothetical protein; 68171 MW; 159EDFDE80EE98D CRC64;
SQ SEQUENCE 615 AA; 68171 MW; 159EDFDE80EE98D CRC64;

Query Match          51.0%; Score 50; DB 4; Length 615;
Best Local Similarity 83.3%; Pred. No. 9.9;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 4 SG1ASGRLYIS 15
Db 454 AGAASGRLYIS 465

RESULT 6
ID 09D7P9 PRELIMINARY; PRT; 423 AA.
AC 09D7P9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE 220003F07R1K protein.
GN 220003F07R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TOES;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arai K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner U., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Segal T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -; SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AK009018; BAB26028.1; -.
DR HSSP: P05619; 1HLE.
DR MGD: MGI:1919119; 2300003F07R1K.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 423 AA; 47834 MW; 5A22BE2FE51B6120 CRC64;

```

```

Query Match          49.0%; Score 48; DB 11; Length 423;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Oy 1 ADLSGASGRLYISRMKG 20
Db 346 ADLTKISKNLYLSKIVHK 365

RESULT 7
ID 09RM57 PRELIMINARY; PRT; 591 AA.
AC 09RM57;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Serine protease, subtilase family.
GN DR0812.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans RI.";
RL Science 286:1571-1577(1999).
DR EMBL: AF001935; AAF10389.1; -.
DR HSSP: Q99405; IMPT.
DR TIGR: DR0812; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003137; PA.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 2.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
KW Protease; Complete proteome.
SQ SEQUENCE 591 AA; 59158 MW; 2EBAD3440D1FA5E4 CRC64;

Query Match          49.0%; Score 48; DB 16; Length 591;
Best Local Similarity 47.1%; Pred. No. 20;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Oy 3 LSG1ASGRLYISRMKG 19
Db 255 VGVASGVNLYMARVLG 271

RESULT 8
ID 09MZ16 PRELIMINARY; PRT; 327 AA.
AC 09MZ16;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GNRH receptor.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA MEDLINE-20267228; PubMed-10809231;
RX Cui J., Smith R.G., Mount G.R., Lo J.L., Yu J., Walsh T.F.,
RA Singh S.B., Devita R.J., Goulet M.T., Schaeffer J.M., Cheng K.;
RT "Identification of Phe313 of the gonadotropin-releasing hormone (GNRH)
RT receptor as a site critical for the binding of nonpeptide GnRH
RT antagonists."
RL Mol. Endocrinol. 14:671-681(2000).
DR EMBL; AF206513; AAF87097.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN.1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
SQ SEQUENCE 327 AA; 37655 MW; 778A82A9D25D3371 CRC64;

Query Match 46.9%; Score 46; DB 6; Length 327;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 LSGIAGSGRLYISRM 18
DB 165 LSGIFAGPOLYIFRM 180
|||||:|:|||||:|

RESULT 9
ID 008804 PRELIMINARY; PRT; 377 AA.
AC 008804;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine proteinase inhibitor NK13.
GN SP112.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL;
RA MEDLINE-97326124; PubMed-9182575;
RX Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;
RT "A new family of 10 murine ovalbumin serpins includes two homologs of
RT proteinase inhibitor 8 and two homologs of the granzyme B inhibitor
RT (proteinase inhibitor 9).";
RL J. Biol. Chem. 272:15434-15441(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL;
RA Kalsnerman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
RA McKeone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
RA Ragoussis J., Bird P.I.;
RT "Comparison of human serpin gene 6p25 with murine chromosome 13 reveals
RT a greatly expanded Ov-serpin gene repertoire in the mouse.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; U96707; AAB57819.2;
DR MGD; MGI:894688; Sp112.
DR InterPro: IPR000215; Serpin.
DR Pfam; PF00079; serpin.1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
DR Serpin.
KW SEQUENCE 377 AA; 42535 MW; 9660B8F3A41EE47A CRC64;

Query Match 46.9%; Score 46; DB 11; Length 377;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGIAGSGRLYISRM 20
DB 11 ADLSGIAGSGRLYISRM 20
|||||:|:|||||:|

```

```

DB 302 ADFSGIAGSGRLYISRM 321

RESULT 10
ID 0967L9 PRELIMINARY; PRT; 382 AA.
AC 0967L9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serpin (Fragment).
GN SPI.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHINESE;
RA Yan Y.T., Liu S.X.;
RT "Serpins Sequence of Schistosoma japonicum.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317731; AAC45932.1;
DR InterPro: IPR000215; Serpin.
DR Pfam; PF00079; serpin.1.
DR SMART; SM00093; SERPIN; 1.
DR NON_TER 1
FT 1
SQ SEQUENCE 382 AA; 43162 MW; B6B816CFA668D1E7 CRC64;

Query Match 46.9%; Score 46; DB 5; Length 382;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSGIAGSGRLYIS 15
DB 302 ADLSGIAGSGRLYIS 316
|||||:|:|||||:|

RESULT 11
ID 0967L9 PRELIMINARY; PRT; 400 AA.
AC 0967L9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Serine proteinase inhibitor serpin.
GN SPI.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHINESE;
RA Yan Y.T., Liu S.X., Song G.C.;
RT "Full-length of Serpin Sequence of Schistosoma japonicum.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF380366; AAK57435.1;
DR InterPro: IPR000215; Serpin.
DR Pfam; PF00079; serpin.1.
DR Protease; Serpin.
KW SEQUENCE 400 AA; 45207 MW; 36BB918B393689BF CRC64;

Query Match 46.9%; Score 46; DB 5; Length 400;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ADLSGIAGSGRLYIS 15
DB 320 ADLSGIAGSGRLYIS 334
|||||:|:|||||:|

```

RESULT 12

09NH65
ID 09NH65 PRELIMINARY; PRT; 373 AA.
AC 09NH65;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Serine proteinase inhibitor.
OS Trichinella spiralis.
OC Eukaryota; Metazoa; Nematoda; Enoplaea; Trichocephalida;
OC Trichinelidae; Trichinella.
OX NCBI_TaxID=6334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISS413;
RA Nagano I., Wu Z., Takehashi Y.;
RT "Molecular cloning and expression of serine proteinase inhibitor from Trichinella spiralis."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AF231948; AAF63473.1; -
DR HSSP: P01008; IATH.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; Serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN.1.
KM Serpin.
SQ SEQUENCE 373 AA; 42414 MW; 78CAACCDAC06BCF5 CRC64;
Query Match 45.98; Score 45; DB 5; Length 373;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
Db 1 ADLSTGASGRLYISRMGK 20
293 ADLSTGCKEKLVSIVHK 312
RESULT 13
ID 026509 PRELIMINARY; PRT; 385 AA.
AC 026509;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE Serpin (Fragment).
GN SPI.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RA Ogunlunde J.O., Raashid A.A., Licate L.S., Liu S., Blanton R.E.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 280-385 FROM N.A.
RC STRAIN=PHILIPPINE;
RA Hooker C.W., Brindley P.J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: U11023; AAA50230.1; -
DR EMBL: AF043997; AAC02994.1; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; Serpin.1.
DR SMART: SM00093; SERPIN.1.
KM Serpin.
FT NON_TER
SQ SEQUENCE 385 AA; 43439 MW; 652E15AD670C4719 CRC64;
Query Match 45.98; Score 45; DB 5; Length 385;
Best Local Similarity 64.3%; Pred. No. 40;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ADLSTGASGRLYI 14
Db 305 ADLSTGSKNNLV 318
RESULT 14
ID 065309 PRELIMINARY; PRT; 942 AA.
AC 065309;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 104.5 kDa protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. WS;
RA Sanders P.M., Bul A.O., Weterings K., McIntire K.N., Hsu Y.C.,
RA Lee P.Y., Truong M.T., Beals T.B., Goldberg R.B.;
RT "Another Development Defects in Arabidopsis thaliana Male-Sterile Mutants."
RL Sex. Plant Reprod. 11:297-322(1999).
DR EMBL: AF060248; AAC97105.1; -
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000051; SAM_bind.
DR Pfam: PF00035; dsrm.1.
KM Hypothetical protein.
SQ SEQUENCE 942 AA; 104497 MW; E09B964161CABF69 CRC64;
Query Match 45.98; Score 45; DB 10; Length 942;
Best Local Similarity 45.0%; Pred. No. 11e+02;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Db 250 AERLGLKDGQVMISRFK 269
RESULT 15
ID 09C508 PRELIMINARY; PRT; 942 AA.
AC 09C508;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CORYMBOSA2.
GN CRM2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Suzuki M., Takahashi T., Kameda Y.;
RL "CORYMBOSA2 (CRM2), a gene required for inflorescence development in Arabidopsis thaliana."
RT Arabidopsis thaliana.
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF327068; AKI16435.1; -
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000051; SAM_bind.
DR Pfam: PF00035; dsrm.1.
SQ SEQUENCE 942 AA; 104454 MW; 7DBB5222C0D28779 CRC64;
Query Match 45.98; Score 45; DB 10; Length 942;
Best Local Similarity 45.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Qy 1 ADLSTGASGRLYISRMGK 20

db |: |: |: |||| |
 250 AERIGKDGNOVMISRMFGK 269

Search completed: July 11, 2003, 11:59:02
Job time : 10.7318 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 2.86902 Seconds
(without alignments)
232.223 Million cell updates/sec

Title: US-10-091-442-7

Perfect score: 26

Sequence: 1 LYDAK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	5	AAK57113	Human megakaryocyte
2	26	100.0	58	AA16557	Bacteriophage 44AH
3	26	100.0	58	ABH81848	Staphylococcus aur
4	26	100.0	69	ABP29986	Streptococcus poly
5	26	100.0	70	ABP29293	Streptococcus poly
6	26	100.0	118	AAK51492	Human secreted pro
7	26	100.0	118	AAK51494	Human secreted pro
8	26	100.0	151	AAK28475	Arabidopsis thalia
9	26	100.0	155	AAK28474	Arabidopsis thalia
10	26	100.0	212	ABG24830	Novel human diagno

11	26	100.0	253	23	ABB49942
12	26	100.0	263	21	AAK25581
13	26	100.0	263	21	AAK37820
14	26	100.0	297	21	AAK25580
15	26	100.0	297	21	AAK37819
16	26	100.0	319	22	AAK79071
17	26	100.0	357	21	AAK25579
18	26	100.0	357	21	AAK37818
19	26	100.0	357	23	ABK93723
20	26	100.0	380	15	AAK48379
21	26	100.0	380	21	AAK08254
22	26	100.0	380	21	AAK24142
23	26	100.0	380	22	AAK83075
24	26	100.0	413	22	AAK90403
25	26	100.0	427	22	AAK79070
26	26	100.0	491	22	AAK10377
27	26	100.0	588	22	AAK15599
28	26	100.0	630	21	AAK57586
29	26	100.0	745	22	AAK36616
30	26	100.0	969	22	AAK71685
31	26	100.0	1014	22	AAK90902
32	26	100.0	1043	18	AAK18306
33	26	100.0	1043	19	AAK56569
34	26	100.0	1213	22	AAK71439
35	26	100.0	1585	22	ABG06075
36	26	100.0	1585	22	ABG10815
37	26	100.0	1598	22	ABG06618
38	26	100.0	1647	22	ABG09655
39	26	100.0	1647	22	ABG04001
40	26	100.0	1844	21	AAK18250
41	26	100.0	2030	22	ABK64300
42	26	100.0	2424	22	ABK68256
43	26	100.0	2918	22	ABG27218
44	26	100.0	73	22	AAK5183
45	24	92.3	88	19	AAK77753

ALIGNMENTS

RESULT 1	AAK57113	standard: peptide; 5 AA.
ID	AAK57113	
XX	AAK57113;	
AC	AAK57113;	
XX		
DT	16-AUG-1994 (first entry)	
XX		
DE	Human megakaryocyte differentiation factor peptide 7.	
XX		
KW	Human megakaryocyte differentiation factor; MDF; thrombopoietin;	
KW	haematopoietic stimulating factor; thrombocytopoenta; platelet;	
KW	bone marrow transplantation; cancer chemotherapy.	
XX		
OS	Homo sapiens.	
XX		
PN	EP583884-A.	
XX		
PD	23-FEB-1994.	
XX		
PF	19-JUL-1993; 93EP-0305654.	
XX		
PR	17-JUL-1992; 92JP-0212305.	
XX		
PR	04-MAR-1993; 93JP-0067339.	
XX		
PA	(SUNR) SUNTORY LTD.	
XX	(TSUTS) TSUTIMOTO M.	
PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;	
XX	Tsujimoto M, Tsunoka N, Yamaguchi N, Yamachi K;	
DR	WPI; 1994-058782/08.	
XX		

Listeria monocytog
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Corynebacterium g1
Arabidopsis thalia
Arabidopsis thalia
Hebically activ
Human megakaryocyt
Human megas in prote
Human megas in prote
C glutamicum prote
Corynebacterium g1
Salmonella typhi (
Novel human diagno
Human papillomavir
Staphylococcus aur
Drosophila melanog
C glutamicum prote
Photobacterium lum
Toxin TccC, encode
Drosophila melanog
Novel human diagno
Novel human diagno
Novel human diagno
Plasmodium falci
Drosophila melanog
Novel human diagno
Human immune/haema
pyruvate oxidase p

PT New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PS decrease in platelets
XX
PS Claim 1; Page 20; 47pp; English.
XX
CC Human MDF (see AAR48379) can be isolated from a culture of human
CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
CC stimulates differentiation of megakaryocytes from myeloid cells
CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
CC making it useful for treatment of diseases involving a decrease
CC in platelet number (esp. thrombocytopenia) such as occurs in bone
CC marrow transplantation and in chemotherapy. MDF has mol.wt.
CC 55-57kD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
CC contains an amino acid sequence comprising at least one of the
CC sequences AAR57107-R57115.
XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 26; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 1 LYDAK 5

RESULT 2
AAB16557
ID AAB16557 standard; Protein; 58 AA.
XX
AC AAB16557;
XX
DT 27-OCT-2000 (first entry)
XX
DE Bacteriophage 44AHJD protein sequence 44AHJDF025.
XX
KW Bacteriophage; antimicrobial; genome; identification; antibacterial;
KW bacterial growth inhibition; bacterial infection.
XX
OS Bacteriophage 44AHJD.
XX
PN WO200032825-A2.
XX
PD 08-JUN-2000.
XX
PF 03-DEC-1999; 99WO-1B02040.
XX
PR 03-DEC-1998; 98US-0110992.
PR 03-JUN-1999; 99US-0326144.
PR 28-SEP-1999; 99US-0407804.
PR 30-SEP-1999; 99US-0157218.
PR 01-DEC-1998; 99US-0168777.
PR 02-DEC-1999; 99US-0434252.
XX
PA (PHAG-) PHAGETECH INC.
XX
PI Pelletier J, Gros P, Dubow M;
XX
DR WPI; 2000-412361/35.
DR N-PSDB; AAA69042.
XX
PT Identifying a bacteriophage coding region for treating bacterial
PT infections comprises identifying a nucleic acid encoding a product that
PT inhibits bacteria when a bacteriophage infects a bacterium
XX
PS Example 9; Page 278; 456pp; English.
XX
CC The present invention describes a method for identifying a bacteriophage
CC coding region encoding a product active on an essential bacterial
CC target. The method comprises identifying a nucleic acid sequence encoding
CC a gene product that provides a bacteria-inhibiting function when an

CC uncharacterised bacteriophage infects a pathogenic bacterium. The
CC compound active on a target of a bacteriophage inhibitor protein in a
CC bacteria is used to treat or prevent a bacterial infection in an animal.
CC AAA68243 to AAA69442 and AAB16523 to AAB16954 represent bacteriophage
CC nucleotide and protein sequences which are used in the exemplification of
CC the present invention.
XX
SQ Sequence 58 AA;

Query Match 100.0%; Score 26; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 29 LYDAK 33

RESULT 3
ABB81848
ID ABB81848 standard; Protein; 58 AA.
XX
AC ABB81848;
XX
DT 20-SEP-2002 (first entry)
XX
DE Staphylococcus aureus 44AHJD ORF 25 amino acid sequence.
XX
KW STAU-R2; drug screening; antibacterial; bacterial infection;
KW 44AHJD ORF 25.
XX
OS Staphylococcus aureus.
XX
PN WO200244718-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-CA01754.
XX
PR 30-NOV-2000; 2000US-0727892.
PR 20-JUN-2001; 2001US-0885561.
XX
PA (PHAG-) PHAGETECH INC.
XX
PI Pelletier J, Gros P, Dubow M;
XX
DR WPI; 2002-566592/60.
DR N-PSDB; ABB86788.
XX
PT Use of a Staphylococcus aureus (STAU-R2) polypeptide or composition
PT comprising STAU-R2 and polypeptide encoded by bacteriophage open
PT reading frame that specifically interacts with STAU-R2, for identifying
PT a compound active on STAU-R2
XX
PS Claim 2; Fig 2; 131pp; English.
XX
CC The invention relates to a novel staphylococcus aureus polypeptide
CC (STAU-R2), and it's use in drug screening assays. The polypeptide of the
CC invention has antibacterial activity. The compound active on STAU-R2 is
CC useful for the manufacture of an antibacterial agent or for the
CC manufacture for medicament for treating or preventing a bacterial
CC infection. The sequence represents the amino acid sequence of S. aureus
CC 44AHJD ORF 25.
XX
SQ Sequence 58 AA;

Query Match 100.0%; Score 26; DB 23; Length 58;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 29 LYDAK 33

RESULT 4

ABP29986
ID ABP29986 standard; Protein: 69 AA.

XX AC ABP29986;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 9148.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;

XX KW antiInflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus pyogenes.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX PI Tettelin H;

XX DR WPI: 2002-352536/38.

XX DR N-PSDB; ABN70617.

XX PT New Streptococcus protein for the treatment or prevention of infection

XX PT or disease caused by Streptococcus bacteria, such as meningitis, and

XX PT for detecting a compound that binds to the protein -

XX PS Claim 1; Page 4037; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B

XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX CC the specification. The proteins have antibacterial and antiinflammatory

XX CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX CC antibodies that bind (I) are used in the manufacture of medicaments for

XX CC the treatment or prevention of infection or disease caused by

XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a

XX CC biological sample. (I) is used to determine whether a compound binds to

XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX CC used as a vaccine or diagnostic composition. The disease caused by

XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic

XX CC acid encoding (I) may be used to recombinantly produce (I) and may be

XX CC used in gene therapy. Antibodies to (I) are used for affinity

XX CC chromatography, immunoassays, and distinguishing/identifying

XX CC Streptococcus proteins.

XX SQ Sequence 69 AA;

XX Query Match 100.0%; Score 26; DB 23; Length 69;

XX Best Local Similarity 100.0%; Pred. No. 76;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 LYDAK 5

XX | | | | |

XX DB 32 LYDAK 36

XX RESULT 5

ABP29293
ID ABP29293 standard; Protein: 70 AA.

XX AC ABP29293;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 7762.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;

XX KW antiInflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus pyogenes.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX PI Tettelin H;

XX DR WPI: 2002-352536/38.

XX DR N-PSDB; ABN69924.

XX PT New Streptococcus protein for the treatment or prevention of infection

XX PT or disease caused by Streptococcus bacteria, such as meningitis, and

XX PT for detecting a compound that binds to the protein -

XX PS Claim 1; Page 3912; 4525pp; English.

XX CC The invention relates to a protein (ABP25413-ABP30895) from group B

XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX CC the specification. The proteins have antibacterial and antiinflammatory

XX CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX CC antibodies that bind (I) are used in the manufacture of medicaments for

XX CC the treatment or prevention of infection or disease caused by

XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a

XX CC biological sample. (I) is used to determine whether a compound binds to

XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX CC used as a vaccine or diagnostic composition. The disease caused by

XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic

XX CC acid encoding (I) may be used to recombinantly produce (I) and may be

XX CC used in gene therapy. Antibodies to (I) are used for affinity

XX CC chromatography, immunoassays, and distinguishing/identifying

XX CC Streptococcus proteins.

XX SQ Sequence 70 AA;

XX Query Match 100.0%; Score 26; DB 23; Length 70;

XX Best Local Similarity 100.0%; Pred. No. 77;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 LYDAK 5

XX | | | | |

XX DB 33 LYDAK 37

XX RESULT 6

XX AAB51492

XX ID AAB51492 standard; Protein: 118 AA.

AC	AAB51492;
XX	
DT	16-FEB-2001 (first entry)
XX	
DE	Human secreted protein BLAST search protein SEQ ID NO: 169.
XX	
KW	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX	neurological disease; infection; human; secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO200058495-A1.
PD	05-OCT-2000.
XX	
PF	23-MAR-2000; 2000WO-US07661.
XX	
PR	26-MAR-1999; 99US-0126504.
PR	07-JAN-2000; 2000US-0174847.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruden SM, Komatsoulis G;
DR	WPI; 2000-611720/58.
XX	
PT	New nucleic acid molecules encoding 45 human secreted proteins for
PT	diagnosing, preventing, treating or ameliorating medical conditions and
PT	used as food additives or preservatives -
PS	
XX	Disclosure; Page 400; 410pp; English.
CC	
CC	The invention relates to the isolation of genes AAc93310-C93354 encoding
CC	the human secreted proteins AAB51380-B51423. The genes and proteins are
CC	useful for preventing, ameliorating or treating medical conditions, e.g.
CC	by protein or gene therapy. The genes are isolated from a range of human
CC	tissues disclosed in the specification. The nucleic acids, proteins,
CC	antibodies and (ant)agonists are useful in the diagnosis, treatment and
CC	prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
CC	cancers of the adrenal gland, bone, bone marrow, breast,
CC	gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC	e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC	autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC	sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC	cardiovascular disorders such as myocardial ischaemias; (d) wound
CC	healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC	and (f) infectious diseases such as viral, bacterial, fungal and
CC	parasitic infections. The present sequence is a protein isolated in the
CC	present invention.
XX	
SQ	Sequence 118 AA:
QY	
DB	
Query Match	100.0%; Score 26; DB 21; Length 118;
Best Local Similarity	100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1 LYDAK 5	
77 LYDAK 81	
RESULT 7	
AAB51494	
ID	AAB51494 standard; Protein; 118 AA.
XX	
AC	AAB51494;
XX	
DT	16-FEB-2001 (first entry)
XX	
DE	Human secreted protein BLAST search protein SEQ ID NO: 171.

XX	Cytostatic; immunosuppressive; neutropic; neuroprotective; antiviral;
KM	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM	cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX	neurological disease; infection; human; secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO200058495-A1.
XX	
PD	05-OCT-2000.
XX	
PF	23-MAR-2000; 2000WO-US07661.
XX	
PR	26-MAR-1999; 99US-0126504.
PR	07-JAN-2000; 2000US-0174847.
XX	
PA	(HUMA-) HUMAN GENOME SCL INC.
PI	Rosen CA, Ruben SM, Komatsoulis G;
XX	
DR	WPI, 2000-611720/58.
XX	
PT	New nucleic acid molecules encoding 45 human secreted proteins for
PT	diagnosing, preventing, treating or ameliorating medical conditions and
XX	used as food additives or preservatives -
PS	
XX	Disclosure; Page 401; 410pp; English.
CC	
CC	The invention relates to the isolation of genes AHC93310-C93354 encoding
CC	the human secreted proteins AAB51380-B51423. This sequence represents a
CC	peptide fragment homologous to the protein encoded by the gene isolated
CC	in the present invention. The sequence is a search result from a BLASTX
CC	homology search. The genes and proteins are useful for preventing,
CC	ameliorating or treating medical conditions, e.g. by protein or gene
CC	therapy. The genes are isolated from a range of human tissues disclosed
CC	in the specification. The nucleic acids, proteins, antibodies and
CC	(ant)agonists are useful in the diagnosis, treatment and prevention of:
CC	(a) cancer, e.g. breast and ovarian cancer, and other cancers of the
CC	adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC	lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC	allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC	diabetes mellitus, Crohn's disease, multiple sclerosis, Rheumatoid
CC	arthritis and ulcerative colitis; (c) cardiovascular disorders such
CC	as myocardial ischaemias; (d) wound healing; (e) neurological diseases
CC	e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such
CC	as viral, bacterial, fungal and parasitic infections.
XX	
SO	Sequence 118 AA;
	Query Match 100.0%; Score 26; DB 21; Length 118;
	Best Local Similarity 100.0%; Pred. No. 14e+02;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 LYDAK 5
Db	77 LYDAK 81
RESULT 8	
AAAG28475	
ID	AAAG28475 standard; Protein: 151 AA.
XX	
AC	AAAG28475;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 33705.
KM	Protein identification; signal transduction pathway; metabolic pathway;
KM	hybridisation assay; genetic mapping; gene expression control; promoter;
FW	termination sequence.

Fri Jul 11 15:11:17 2003

us-10-091-442-7.rag

```
XX Arbidopsis thaliana.
OS
XX EPI033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134421.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135153.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
```

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155859.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match Best Local Similarity 100.0%; Score 26; DB 21; Length 151;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
11111
Db 2 LYDAK 6

RESULT 9

AAG28474 ID AAG28474 standard; Protein; 155 AA.

AC AAG28474;

DF 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 33704.

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

XX

PR 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145108.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 26; DB 21; Length 155;
Best Local Similarity 100.0%; Pred. No. 18e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LYPDAK 5
11111
Db 6 LYPDAK 10

RESULT 10
ABG24830
ID ABG24830 standard; protein; 212 AA.
XX
AC ABG24830;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24821.
XX
KW Human; chromosome mapping; gene mapping; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.
 DR N-PSDB; AAS89017.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 55189; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG3037 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 SQ Sequence 212 AA:

Query Match 100.0%; Score 26; DB 22; Length 212;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LYDAR 5
 |||||
 Db 37 LYDAR 41

RESULT 11
 ABB49942
 ID ABB49942 standard; Protein; 253 AA.
 XX
 AC ABB49942;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #2646.
 XX
 DE Listeria monocytogenes protein #2646.
 XX
 KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 PN WO200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX
 PR 11-APR-2000; 2000FR-0004629.
 XX
 PA (INSP) INST PASTEUR;
 PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P,
 PI Dusurget O, Chetoui F, Nedjati H, Glaser P, Kunst F, Cossart P,
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Madueno E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 DR WPI: 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides -
 PS Claim 6; SEQ ID No 2647; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies. Identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 SQ Sequence 253 AA:

Query Match 100.0%; Score 26; DB 23; Length 253;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LYDAR 5
 |||||
 Db 231 LYDAR 235

RESULT 12
 AAG25581
 ID AAG25581 standard; Protein; 263 AA.
 XX
 AC AAG25581;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 29703.
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 29703.
 XX
 KM Protein identification; signal transduction pathway; metabolic pathway;
 KM hybridisation assay; genetic mapping; gene expression control; promoter;
 KM termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 29-MAR-1999; 99US-0126264.
 PR 01-APR-1999; 99US-0126785.
 PR 06-APR-1999; 99US-0127462.
 PR 99US-0128234.

PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 23-APR-1999: 99US-0130891.
PR 28-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 04-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 07-MAY-1999: 99US-0132487.
PR 11-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134421.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142820.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 13-JUL-1999: 99US-0143624.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.

PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144332.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144684.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145913.
PR 27-JUL-1999: 99US-0145918.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0146389.
PR 04-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 05-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 06-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 09-AUG-1999: 99US-0147935.
PR 10-AUG-1999: 99US-0148171.
PR 11-AUG-1999: 99US-0148319.
PR 12-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 13-AUG-1999: 99US-0148584.
PR 16-AUG-1999: 99US-0149368.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155486.
PR 28-SEP-1999: 99US-0155659.
PR 29-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 26; DB 21; Length 263;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 114 LYDAK 118

RESULT 14
AAC25580
ID AAC25580 standard; Protein; 297 AA.
XX
AC AAG25580;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29702.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134321.
 PR 14-MAY-1999; 99US-0134320.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149923.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0158294.
 PR 13-OCT-1999; 99US-0158295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

Query Match 100.0%; Score 26; DB 21; Length 297;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LYDAK 5
Db 148 LYDAK 152

RESULT 15
AAG37819
ID AAG37819 standard; Protein: 297 AA.
XX AAG37819:
AC
XX
XX
DT 18-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46565.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127462.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 23-APR-1999: 99US-0130891.
PR 28-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.
PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134221.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.

PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144844.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 22-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145313.
PR 27-JUL-1999: 99US-0145318.
PR 27-JUL-1999: 99US-0145319.
PR 28-JUL-1999: 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 26; DB 21; Length 297;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 DB 148 LYDAK 152

Search completed: July 11, 2003, 11:54:24
 Job time : 4.86902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: July 11, 2003, 11:49:36 ; Search time 0.893971 seconds
(Without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-7

Perfect score: 26

Sequence: 1 LVDK 5

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfilsl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	5	2	US-08-472-659-7
2	26	100.0	5	2	US-08-474-661-7
3	26	100.0	5	2	US-08-611-977-7
4	26	100.0	380	2	US-08-472-659-34
5	26	100.0	380	2	US-08-474-661-34
6	26	100.0	380	2	US-08-611-977-34
7	26	100.0	630	4	US-09-300-909-20
8	24	92.3	88	4	US-08-936-165A-513
9	24	92.3	377	3	US-09-023-023-1
10	24	92.3	388	2	US-08-742-621-1
11	24	92.3	388	4	US-09-191-608-22
12	24	92.3	471	1	US-08-203-905B-2
13	24	92.3	472	1	US-08-203-905B-14
14	24	92.3	586	4	US-09-134-001C-4456
15	23	88.5	11	4	US-09-177-249-236
16	23	88.5	75	4	US-08-858-207A-460
17	23	88.5	77	4	US-08-803-346-64
18	23	88.5	155	4	US-09-129-030-6
19	23	88.5	178	4	US-09-129-030-54
20	23	88.5	185	4	US-09-387-800-2
21	23	88.5	284	2	US-08-751-233A-6
22	23	88.5	284	2	US-08-751-233A-8
23	23	88.5	308	4	US-09-323-872A-34
24	23	88.5	434	4	US-08-370-190-3
25	23	88.5	434	4	US-09-236-615-3
26	23	88.5	436	2	US-08-537-400-3
27	23	88.5	452	2	US-08-731-079A-2

28	23	88.5	547	4	US-09-178-252-11	Sequence 11, Appl
29	23	88.5	569	1	US-08-306-231-3	Sequence 3, Appl
30	23	88.5	585	2	US-08-426-125-2	Sequence 2, Appl
31	23	88.5	585	2	US-08-455-355-2	Sequence 4, Appl
32	23	88.5	876	1	US-08-717-515-4	Sequence 2, Appl
33	23	88.5	893	2	US-08-706-702-3	Sequence 3, Appl
34	23	88.5	893	3	US-08-484-661A-2	Sequence 3, Appl
35	23	88.5	893	3	US-08-706-706-3	Sequence 2, Appl
36	23	88.5	893	4	US-09-019-160-2	Sequence 6, Appl
37	23	88.5	893	4	US-09-019-160-6	Sequence 7, Appl
38	23	88.5	893	4	US-09-019-160-7	Sequence 8, Appl
39	23	88.5	893	4	US-09-019-160-8	Sequence 9, Appl
40	23	88.5	893	4	US-09-019-160-9	Sequence 2, Appl
41	23	88.5	893	5	PCT-US96-09641-2	Sequence 10, Appl
42	23	88.5	896	2	US-08-640-389A-10	Sequence 12, Appl
43	23	88.5	896	2	US-08-640-389A-12	Sequence 10, Appl
44	23	88.5	896	4	US-08-618-957A-10	
45	23	88.5				

ALIGNMENTS

RESULT 1
US-08-472-659-7
Sequence 7, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUDIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROJOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhitro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030com1
TITLE OF INVENTION: MECAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-472-659-7

Query Match 100.0%; Score 26; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
11111
DB 1 LYDAK 5

RESULT 2
US-08-474-661-7
Sequence 7, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5874253uo

APPLICANT: NAKAZATO, Hitoshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5874253uhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAGUCHI, Kozo

APPLICANT: YAMAGUCHI, No. 5874253omi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474, 661

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEX

REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 001560-204

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-661-7

Query Match 100.0%; Score 26; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
11111
DB 1 LYDAK 5

RESULT 3
US-08-611-977-7
Sequence 7, Application US/08611977
Patent No. 5972886

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5972886uo

APPLICANT: NAKAZATO, Hitoshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5972886uhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAGUCHI, Kozo

APPLICANT: YAMAGUCHI, No. 5972886omi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611, 977

FILING DATE: 06-MAR-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-204

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-7

Query Match 100.0%; Score 26; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
11111
DB 1 LYDAK 5

RESULT 4
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 26; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 121 LYDAK 125

RESULT 5
US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki

APPLICANT: TSUROOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 26; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 121 LYDAK 125

RESULT 6
US-08-611-977-34
Sequence 34, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22113-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/611,977
;; FILING DATE: 06-MAR-1996
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MCGOWAN, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-204
;;
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;;
;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 380 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-611-977-34

Query Match
Best Local Similarity 100.0%; Score 26; DB 2; Length 380;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 121 LYDAK 125

RESULT 7
US-09-300-909-20
; Sequence 20, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING
; NUMBER OF SEQUENCES: 27
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,909
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/083,942
; FILING DATE: 01-MAY-1998
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 amino acids
; TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-09-300-909-20

Query Match
Best Local Similarity 100.0%; Score 26; DB 4; Length 630;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 493 LYDAK 497

RESULT 8
US-08-936-165A-513
; Sequence 513, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582e1 Prokaryotic polynucleotides,
; NUMBER OF SEQUENCES: 534
; NUMBER OF SEQUENCES: 534
; INFORMATION FOR SEQ ID NO: 513:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-936-165A-513

Query Match
Best Local Similarity 92.3%; Score 24; DB 4; Length 88;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5

Db :||||
48 MYDAK 52

RESULT 9
US-09-023-023-2
Sequence 2, Application US/09023023
Patent No. 6121018
GENERAL INFORMATION:
APPLICANT: Kristine Kay Kikly
TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme Like Apoptosis Pr
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,023
FILING DATE: 12-FEB-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/042,030
FILING DATE: March 27, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GH50013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-023-023-2

Query Match
Best Local Similarity 92.3%; Score 24; DB 3; Length 377;
Pred. No. 5.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 45 IYDAK 49

RESULT 10
US-08-742-621-1
Sequence 1, Application US/08742621
Patent No. 5856129
GENERAL INFORMATION:
APPLICANT: HILLMAN, JENNIFER L.
TITLE OF INVENTION: NOVEL HUMAN PURINOCEPTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,621
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0147 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: consensus
US-08-742-621-1

Query Match
Best Local Similarity 92.3%; Score 24; DB 2; Length 388;
Pred. No. 5.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 218 IYDAK 222

RESULT 11
US-09-191-608-22
Sequence 22, Application US/09191608
Patent No. 6242216
GENERAL INFORMATION:
APPLICANT: Lynch, Kevin J.
APPLICANT: Burgard, Edward C.
APPLICANT: Metzger, Randy E.
APPLICANT: Niforatos, Wende
APPLICANT: Touma, Edward B.
APPLICANT: Van Biesen, T.
TITLE OF INVENTION: Nucleic Acids Encoding a Functional
TITLE OF INVENTION: Human Purinoceptor P2X2 and P2X4 and Methods of Production
FILE REFERENCE: 6394.US.P1
CURRENT APPLICATION NUMBER: US/09/191,608
CURRENT FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 22
LENGTH: 388
TYPE: PRT
ORGANISM: Homo sapiens
US-09-191-608-22

Query Match
Best Local Similarity 92.3%; Score 24; DB 4; Length 388;
Pred. No. 5.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 218 IYDAK 222

RESULT 12

US-08-203-905B-2
Sequence 2, Application US/08203905B
Patent No. 5646249
GENERAL INFORMATION:
APPLICANT: KAYE, FEDERIC J.
APPLICANT: OTTERSON, GREGORY A.
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A
TITLE OF INVENTION: NOVEL CHAPERONE PROTEIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR
STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/203,905B
FILING DATE: February 28, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: KIRKPATRICK, ANITA M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH089.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-0176
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 471 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-203-905B-2

Query Match
Best Local Similarity 92.3%; Score 24; DB 1; Length 471;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 96 IYDAK 100

RESULT 13
US-08-203-905B-14
Sequence 14, Application US/08203905B
Patent No. 5646249
GENERAL INFORMATION:
APPLICANT: KAYE, FEDERIC J.
APPLICANT: OTTERSON, GREGORY A.
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A
TITLE OF INVENTION: NOVEL CHAPERONE PROTEIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR
STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/203,905B

FILING DATE: February 28, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: KIRKPATRICK, ANITA M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH089.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-0176
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-203-905B-14

Query Match
Best Local Similarity 92.3%; Score 24; DB 1; Length 472;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 97 IYDAK 101

RESULT 14
US-09-134-001C-4456
Sequence 4456, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4456
LENGTH: 586
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4456

Query Match
Best Local Similarity 92.3%; Score 24; DB 4; Length 586;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 94 IYDAK 98

RESULT 15
US-09-177-249-236
Sequence 236, Application US/09177249
Patent No. 6229064
GENERAL INFORMATION:
APPLICANT: Fischer, Robert L.
APPLICANT: Ohad, Nir
APPLICANT: Kiyosue, Tomohiro
APPLICANT: Yadegaril, Ramln
APPLICANT: Margossian, Linda
APPLICANT: Harada, John
APPLICANT: Goldberg, Robert B.
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
FILE REFERENCE: 023070-08612005

; CURRENT APPLICATION NUMBER: US/09/177,249
 ; CURRENT FILING DATE: 1998-10-22
 ; EARLIER APPLICATION NUMBER: US 09/071,838
 ; EARLIER FILING DATE: 1998-05-01
 ; NUMBER OF SEQ ID NOS: 324
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 236
 ; LENGTH: 11
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis sp.
 US-09-177-249-236

Query Match 88.5%; Score 23; DB 4; Length 11;
 Best Local Similarity 80.0%; Pred. No. 22;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LYDAK 5
 |||:|
 Db 1 LYDSK 5

Search completed: July 11, 2003, 12:02:10
 Job time : 1.89397 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 1.4657 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-7

Perfect score: 26

Sequence: 1 LYDAK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications-AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	5	9	US-10-091-442-7
2	26	100.0	5	10	US-09-140-719-7
3	26	100.0	380	9	US-10-091-442-34
4	26	100.0	380	10	US-09-140-719-34
5	26	100.0	413	9	US-09-738-626-4157
6	26	100.0	430	9	US-10-156-761-14800
7	26	100.0	464	9	US-10-036-9598-9
8	26	100.0	491	10	US-09-794-098-2
9	26	100.0	567	9	US-10-156-761-7781
10	26	100.0	707	9	US-10-156-761-7899
11	26	100.0	712	9	US-10-156-761-15026
12	26	100.0	745	10	US-09-815-242-12209
13	26	100.0	1014	9	US-09-738-626-4656
14	26	100.0	1043	9	US-10-242-056-61
15	24	92.3	88	10	US-09-939-980-513
16	24	92.3	91	9	US-10-156-733-10
17	24	92.3	140	12	US-10-010-901-27
18	24	92.3	347	10	US-09-816-028A-17
19	24	92.3	388	10	US-09-833-082-2

20	24	92.3	447	9	US-10-059-585-51	Sequence 51, Appl
21	24	92.3	471	9	US-10-171-311-212	Sequence 212, App
22	24	92.3	579	10	US-09-815-242-5821	Sequence 5821, Ap
23	24	92.3	579	10	US-09-815-242-12943	Sequence 12943, A
24	24	92.3	645	9	US-10-156-761-12752	Sequence 12752, A
25	24	92.3	967	9	US-09-563-728A-30	Sequence 30, Appl
26	24	92.3	967	10	US-09-817-913-7	Sequence 7, Appl1
27	24	92.3	967	10	US-09-817-538-7	Sequence 7, Appl1
28	24	92.3	1084	9	US-10-072-094-7	Sequence 7, Appl1
29	24	92.3	1084	9	US-10-173-539-12	Sequence 12, Appl
30	23	88.5	11	9	US-10-213-512-236	Sequence 236, App
31	23	88.5	11	10	US-09-071-838-236	Sequence 236, App
32	23	88.5	185	9	US-10-234-329-15	Sequence 15, Appl
33	23	88.5	434	9	US-10-112-039-3	Sequence 3, Appl1
34	23	88.5	442	10	US-09-815-242-13237	Sequence 13237, A
35	23	88.5	442	10	US-09-815-242-13702	Sequence 13702, A
36	23	88.5	444	9	US-10-108-605-187	Sequence 187, App
37	23	88.5	448	10	US-09-815-242-13272	Sequence 13272, A
38	23	88.5	454	10	US-09-815-242-10700	Sequence 10700, A
39	23	88.5	498	9	US-10-083-357-1290	Sequence 1290, Ap
40	23	88.5	518	10	US-09-816-248-17	Sequence 17, Appl
41	23	88.5	547	10	US-09-826-660-11	Sequence 11, Appl
42	23	88.5	590	10	US-09-815-242-5099	Sequence 5099, Ap
43	23	88.5	616	9	US-10-126-279-32	Sequence 32, Appl
44	23	88.5	617	9	US-10-126-279-31	Sequence 31, Appl
45	23	88.5	634	10	US-09-816-248-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-091-442-7
Sequence 7, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
INASA, Fuyuki
TSURUOKA, No. US20020164711A1uo
NAKAZATO, Hiroshi
MURA, Kenju
ISHIDA, No. US20020164711A1uhfro
KURIHARA, Tatsuya
YAMAIICHI, Kozo
YAMAGUCHI, No. US20020164711A1om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSER: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: MCGOWAN, MALCOLM K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021.
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 US-10-091-442-7

```

: REGISTRATION NUMBER: 39, 300
: .REFERENCE/DOCKET NUMBER: 001560-247
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID.NO: 7:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 5 amino acids
:   TYPE: amino acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-09-140-719-7

Query Match          100.0%   Score 26;  DB 10;   Length 5;
Best Local Similarity 100.0%   Pred. NO. 4e+05;
Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

```

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 26; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 121 LYDAK 125

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: INASA, Fuyuki
APPLICANT: TSUBOUOKA, No. US20010026931A1uo
APPLICANT: MAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uh1ro
APPLICANT: KURIHARA, Tetsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1om1
TITLE OF INVENTION: MECAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 26; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 121 LYDAK 125

RESULT 5
US-09-738-626-4157
Sequence 4157, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 413
LENGTH: 413
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4157

Query Match 100.0%; Score 26; DB 9; Length 413;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 340 LYDAK 344

RESULT 6
US-10-156-761-14800
Sequence 14800, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30

;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 14800
;; LENGTH: 430
;; TYPE: PRN
;; ORGANISM: Streptomyces avermitilis
US-10-156-761-14800

Query Match
Best Local Similarity 100.0%; Score 26; DB 9; Length 430;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 242 LYDAK 246

RESULT 7
US-10-036-959B-9
;; Sequence 9, Application US/10036959B
;; Publication No. US20030119098A1
;; GENERAL INFORMATION:
;; APPLICANT: E.I. du Pont de Nemours & Company
;; APPLICANT: Hallahan, David L.
;; APPLICANT: Keiper-Hrynko, Natalie
;; TITLE OF INVENTION: Genes involved in the biosynthesis of isopentenyl diphosphate in
;; FILE REFERENCE: CI-1792
;; CURRENT APPLICATION NUMBER: US/10/036,959B
;; CURRENT FILING DATE: 2002-05-10
;; PRIOR APPLICATION NUMBER: 60/307,637
;; PRIOR FILING DATE: 2001-07-25
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 9
;; LENGTH: 464
;; TYPE: PRN
;; ORGANISM: Hevea brasiliensis
US-10-036-959B-9

Query Match
Best Local Similarity 100.0%; Score 26; DB 9; Length 464;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 309 LYDAK 313

RESULT 8
US-09-794-098-2
;; Sequence 2, Application US/09794098
;; Patent No. US20020012668A1
;; GENERAL INFORMATION:
;; APPLICANT: Universiti Sains Malaysia
;; TITLE OF INVENTION: DNA SEQUENCE ENCODING THE SPECIFIC AND ANTIGENIC OUTER MEMBRANE
;; FILE REFERENCE: 4463-005
;; CURRENT APPLICATION NUMBER: US/09/794,098
;; CURRENT FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: P1 2000 0765
;; PRIOR FILING DATE: 2000-02-28
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: Patent version 3.0
;; SEQ ID NO 2
;; LENGTH: 491
;; TYPE: PRN
;; ORGANISM: Salmonella typhi
US-09-794-098-2

Query Match
Best Local Similarity 100.0%; Score 26; DB 10; Length 491;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LYDAK 5
DB 400 LYDAK 404

RESULT 9
US-10-156-761-7781
;; Sequence 7781, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OMURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIRA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRA
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 7781
;; LENGTH: 567
;; TYPE: PRN
;; ORGANISM: Streptomyces avermitilis
US-10-156-761-7781

Query Match
Best Local Similarity 100.0%; Score 26; DB 9; Length 567;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 447 LYDAK 451

RESULT 10
US-10-156-761-7899
;; Sequence 7899, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OMURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIRA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRA
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 7899
;; LENGTH: 707
;; TYPE: PRN
;; ORGANISM: Streptomyces avermitilis
US-10-156-761-7899

Query Match
Best Local Similarity 100.0%; Score 26; DB 9; Length 707;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 470 LYDAK 474

RESULT 11

US-10-156-761-15026
; Sequence 15026, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: SHIBA, TADATOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 15026
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-15026

Query Match 100.0%; Score 26; DB 9; Length 712;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 472 LYDAK 476

RESULT 12

US-09-815-242-12209
; Sequence 12209, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Keril L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12209
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12209

Query Match 100.0%; Score 26; DB 10; Length 745;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 633 LYDAK 637

RESULT 13

US-09-738-626-4656
; Sequence 4656, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4656
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4656

Query Match 100.0%; Score 26; DB 9; Length 1014;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 730 LYDAK 734

RESULT 14

US-10-242-056-61
; Sequence 61, Application US/10242056
; Publication No. US20030113323A1
; GENERAL INFORMATION:
; APPLICANT: Ensign, Jerald C
; APPLICANT: Bowen, David J
; APPLICANT: Petrell, James
; APPLICANT: Fattig, Raymond
; APPLICANT: Schoonover, Sue
; APPLICANT: Firenze-Constant, Richard
; APPLICANT: Orr, Gregory L
; APPLICANT: Merlo, Donald J

APPLICANT: Roberts, Jean L
APPLICANT: Rocheleau, Thomas A
TITLE OF INVENTION: Insecticidal Protein Toxins from
TITLE OF INVENTION: Photorhabdus
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dowelanco
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: IN
COUNTRY: US
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/242,056
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/063,615
FILING DATE: 18-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/395,497
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/743,699
FILING DATE: 06-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50301E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 1043 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-10-242-056-61

Query Match 100.0%; Score 26; DB 9; Length 1043;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 62 LYDAK 66

RESULT 15
US-09-939-980-513
Sequence 513, Application US/09939980
Patent No. US20020082234A1
GENERAL INFORMATION:
APPLICANT: Black, Michael
Burnham, Martin
Hodgson, John
Knowles, David
Lonetto, Michael

Nicholas, Richard
Piatl, Julie
Reichard, Richard
Rosenberg, Martin
Ward, Judith
TITLE OF INVENTION: No. US20020082234A1 Prokaryotic Polynucleotides,
Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,980
FILING DATE: 27-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/936,165
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 513:
SEQUENCE CHARACTERISTICS:
LENGTH: 88 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 513:
US-09-939-980-513

Query Match 92.3%; Score 24; DB 10; Length 88;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 48 MYDAK 52

Search completed: July 11, 2003, 12:37:34
Job time : 2.4657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 1.00832 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-7
Perfect score: 26
Sequence: 1 LYDAK 5

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	117	2 S41811	Ig kappa chain V r
2	26	100.0	178	2 H70345	conserved hypochet
3	26	100.0	197	2 B71106	hypothetical prote
4	26	100.0	199	2 T24871	hypothetical prote
5	26	100.0	253	2 AD1797	partition protein,
6	26	100.0	253	2 AF1423	partition protein,
7	26	100.0	259	2 G81427	periplasmic protei
8	26	100.0	277	2 S73522	probable lipoprote
9	26	100.0	281	2 T50351	probable methyltra
10	26	100.0	301	2 G97187	UDP-glucose 4-epim
11	26	100.0	304	2 C90453	hypothetical prote
12	26	100.0	326	2 T15194	hypothetical prote
13	26	100.0	341	2 B83298	conserved hypochet
14	26	100.0	356	2 E75266	GGDEF family prote
15	26	100.0	401	2 T50679	CP49 protein - chl
16	26	100.0	408	2 S57148	hypothetical prote
17	26	100.0	421	2 H89916	diaminopimelate de
18	26	100.0	439	2 E98139	hypothetical prote
19	26	100.0	439	2 AF3148	hypothetical prote
20	26	100.0	443	2 A27993	methyl coenzyme M
21	26	100.0	450	2 T50631	CP49 protein - chl
22	26	100.0	474	2 T09688	hydroxymethylgluta
23	26	100.0	475	2 T27811	hypothetical prote
24	26	100.0	491	2 S70190	tolC protein - Sal
25	26	100.0	491	2 AB0890	outer membrane pro
26	26	100.0	512	2 B82433	aerobic glycerol-3
27	26	100.0	526	1 C71081	probable helicase
28	26	100.0	543	2 T50905	proteobacteriophillid
29	26	100.0	549	2 A90253	hypothetical prote

30	26	100.0	552	2 B90476	hypothetical prote
31	26	100.0	572	2 S52694	asparagine synthas
32	26	100.0	574	2 G69769	pyruvate oxidase h
33	26	100.0	577	2 D83473	isocitrate dehydro
34	26	100.0	605	2 G95853	probable pyruvate
35	26	100.0	630	1 W1ML35	E1 protein - human
36	26	100.0	633	2 H95932	probable adenylyl-
37	26	100.0	637	2 S36523	E1 protein - human
38	26	100.0	641	1 Z2ERNQ	adenylyl-sulfate k
39	26	100.0	641	2 E95320	adenylyl-sulfate k
40	26	100.0	662	2 T46228	hypothetical prote
41	26	100.0	730	2 S64998	hypothetical prote
42	26	100.0	734	2 AF2001	hypothetical prote
43	26	100.0	745	2 A89770	hypothetical prote
44	26	100.0	768	2 S76977	pled-4 protein - S
45	26	100.0	815	2 E70021	3-hydroxyacyl-CoA

ALIGNMENTS

RESULT 1
S41811
Ig kappa chain V region L22 - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C/Accession: S41811
R:Huber, C.; Schaeble, K.F.; Huber, E.; Klein, R.; Melndl, A.; Thiebe, R.; Lamm, R.; Eur. J. Immunol. 23, 2868-2875, 1993
A:Title: The V(kappa) genes of the L regions and the repertoire of V(kappa) gene sequ
A:Reference number: S41809, PMID:94039386; PMID:8223863
A:Accession: S41811
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <HUB>
A:Cross-references: EMBL:X72816; NID:9415368; PIDN:CAAS1335.1; PID:94388773
C:Genetics:
A:Introns: 19/1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:38-112/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 26; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. NO. 35;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 70 LYDAK 74
|||||

RESULT 2
H70345
conserved hypothetical protein ag_507 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C/Accession: H70345
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300, PMID:98196666; PMID:9537320
A:Accession: H70345
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Residues: 1-178 <AOF>
A:Molecule type: DNA
A:Cross-references: GB:AE000693; NID:92983148; PIDN:AAC06761.1; PID:92983159; GB:AE00
C:Genetics:
A:Experimental source: strain VFS
A:Gene: ag_507

Query Match 100.0%; Score 26; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. NO. 56;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||

Db 61 LYDAK 65

RESULT 3

B71106

hypothetical protein PH0619 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999

C:Accession: B71106

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: B71106

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-197 <KAM>

A:Cross-references: GB:AP000003; NID:93236130; PIDN:BAA29708.1; PID:di030651; PID:932570

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0619

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 197;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||

Db 62 LYDAK 66

RESULT 4

T24871

hypothetical protein T12G3.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T24871

R:Sims, M.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19945

A:Accession: T24871

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-199 <WIL>

A:Cross-references: EMBL:Z68752; PIDN:CNA92984.1; GSPDB:GN00022; CESP:T12G3.5

A:Experimental source: clone T12G3

C:Genetics:

A:Gene: CESP:T12G3.5

A:Map position: 4

A:introns: 45/1; 166/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T12G3.5

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 199;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||

Db 134 LYDAK 138

RESULT 5

AD1797

Partition protein, Para homolog [Imported] - Listeria innocua (strain Clp11262)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002

C:Accession: AD1797

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.;

Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1797

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <GL>

A:Cross-references: GB:AU592022; PIDN:CAC98148.1; PID:916415464; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: para

C:Superfamily: regulatory protein spo0

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 253;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||

Db 231 LYDAK 235

RESULT 6

AF1423

Partition protein, Para homolog para [Imported] - Listeria monocytogenes (strain EGD-

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002

C:Accession: AF1423

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioec

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.;

Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AF1423

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <GL>

A:Cross-references: GB:NC_003210; PIDN:CAD01004.1; PID:916412291; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: para

C:Superfamily: regulatory protein spo0

Query Match

Best Local Similarity 100.0%; Score 26; DB 2; Length 253;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||

Db 231 LYDAK 235

RESULT 7

G81427

periplasmic protein Cj0111 [Imported] - Campylobacter jejuni (strain NCCTC 11168)

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: G81427

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chl

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: AB1250; MUID:20150912; PMID:10688204

A:Accession: G81427

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-259 <PAR>
A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72595.1; PID:g696760
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0111

Query Match 100.0%; Score 26; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 218 LYDAK 222

RESULT 8
S73522
probable lipoprotein E09_orf277 - Mycoplasma pneumoniae (strain ATCC 29342)

N:Alternate names: hypothetical protein E09_orf277
C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73522
R:Hilbert, H.; Hilbert, H.; Plagens, H.; Parkl, E.; Li, B.C.; Herrmann, R.

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Reference number: S73327; MID:97105865; PMID:8948663
A:Accession: S73522

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-277 <HIM>

A:Cross-references: EMBL:AE000020; GB:U00089; NID:g1673852; PIDN:AB95844.1; PID:g167386
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Gene: SGC3

C:Superfamily: hypothetical protein MG440

Query Match 100.0%; Score 26; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 85 LYDAK 89

RESULT 9
T50351
probable methyltransferase [imported] - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000

C:Accession: T50351
R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

A:Submitted to the EMBL Data Library, February 2000

A:Reference number: 225064

A:Accession: T50351

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-281 <SEES>

A:Cross-references: EMBL:AL157918; PIDN:CAB76043.1; GSPDB:GN00067; SPDB:SPBC21C3.07C
C:Genetics:
A:Gene: SPDB:SPBC21C3.07C

A:Map position: 2

A:Introns: 44/2; 259/2

Query Match 100.0%; Score 26; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||

DB 164 LYDAK 168

RESULT 10
G97187

UDP-glucose 4-epimerase [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: G97187

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L.

J. Daily, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MID:21359325; PMID:21359325

A:Accession: G97187

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-301 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK80290.1; PID:g15025343; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:
A:Gene: CAC2334

C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo

Query Match 100.0%; Score 26; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 86 LYDAK 90

RESULT 11
C90453
hypothetical protein hpce-2 [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: C90453

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Ch

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder

arrett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: C90453

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-304 <KUR>

A:Cross-references: GB:AE006641; NID:g13816109; PIDN:AAK42882.1; GSPDB:GN00155

C:Genetics:
A:Gene: hpce-2

Query Match 100.0%; Score 26; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 16+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 35 LYDAK 39

RESULT 12
T15194

hypothetical protein F55A12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999

C:Accession: T15194

R:Pauley, A.; Gattung, S.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid F55A12.

A:Reference number: 218305

A:Accession: T15194

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-326 <PAU>
A:Cross-references: EMBL:AF003130; NID:92088663; PID:92088670; PIDN:AB54128.1; GSPDB:GN
A:Experimental source: strain Bristol N2; clone F55A12
C:Genetics:
A:Gene: CESP:F55A12.2
A:Map position: 1
A:introns: 39/2; 121/1; 165/3; 241/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F55A12.2

Query Match 100.0%; Score 26; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 130 LYDAK 134

RESULT 13
B83298
conserved hypothetical protein PA2771 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83298
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
; Lory, S.; Olson, M.V.
Nucleotide 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: B83298
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-341 <STO>
A:Cross-references: GB:AE004705; GB:AE004091; NID:99948851; PIDN:AAG06159.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2771

Query Match 100.0%; Score 26; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 327 LYDAK 331

RESULT 14
E75266
GGDEF family protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75266
R:White, O.J.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Mc
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75266
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <WHI>
A:Cross-references: GB:AE002079; GB:AE000513; NID:96460315; PIDN:AAF12040.1; PID:9646032
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2498
A:Map position: 1

Query Match 100.0%; Score 26; DB 2; Length 356;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 344 LYDAK 348

RESULT 15
I50679
CP49 protein - chicken
C:Species: Gallus gallus (chicken)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I50679; S58355
R:Sawada, K.; Agata, S.K.; Eguchi, G.; Quinlan, R.A.; Maisel, H.
Curr. Eye Res. 14, 545-553, 1995
A:Title: The predicted structure of chick lens CP49 and a variant thereof, CP49ins, th
A:Reference number: I50631; MUID:9606015; PMID:7587300
A:Accession: I50679
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-401 <SAP>
A:Cross-references: EMBL:X84806; NID:9953184; PIDN:CAA59271.1; PID:94377562
R:Agata, S.K.; Eguchi, G.; Quinlan, R.A.; Maisel, H.
submitted to the EMBL Data Library, February 1995
A:Description: The predicted structure of chick lens CP49 and a variant thereof, CP49I
A:Reference number: S58351
A:Accession: S58355
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-217, 'A', 219-379, 'H', 381-401 <AGA>
A:Cross-references: EMBL:X84806; NID:9953184; PIDN:CAA59271.1; PID:91334742
C:Genetics:
A:Gene: CP49
C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 26; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 335 LYDAK 339

Search completed: July 11, 2003, 12:00:41
Job time : 3.10832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 0.47817 Seconds

(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-7

Perfect score: 26

Sequence: 1 LYDAK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	100.0	277	Y44A_MYCPN	P75151 mycoplasma
2	26	100.0	380	SPB7_HUMAN	O75635 homo sapien
3	26	100.0	408	YJ95_YEAST	P47160 saccharomyc
4	26	100.0	443	MCRB_METVA	P07956 methanococc
5	26	100.0	491	TOLC_SALEN	O54001 salmonella
6	26	100.0	543	BCHB_RHOGE	O91343 rhodocycus
7	26	100.0	571	ASNI_YEAST	P49089 saccharomyc
8	26	100.0	637	VE1_HPV35	P27220 human papil
9	26	100.0	641	NODO_RHIME	P13442 r nodg blfu
10	26	100.0	1091	NCA1_CHICK	P13590 gallus gall
11	26	100.0	1337	P152_YEAST	P39685 saccharomyc
12	26	100.0	179	YH64_SYNR3	P73042 synecocyst
13	26	100.0	257	Y418_METVA	O57861 methanococc
14	26	100.0	292	XOHB_ECOS7	O85605 escherichia
15	26	100.0	292	XOHB_ECOS7	O46800 escherichia
16	26	100.0	309	PNAD_MOUSE	O64311 mus musculu
17	26	100.0	309	PNAD_MOUSE	O28855 sus scrofa
18	26	100.0	310	DCHS_LACS3	P00862 lactobacilli
19	26	100.0	316	RLA0_PLAFB	O94660 plasmodium
20	26	100.0	358	TRMU_RICPR	O94dmi rickettsia
21	26	100.0	369	V181_FOMPV	O95152 fowlpox vir
22	26	100.0	370	TRMU_RICCN	O92110 rickettsia
23	26	100.0	377	ICED_BOVIN	O75601 bos taurus
24	26	100.0	388	P2X4_HUMAN	O95971 homo sapien
25	26	100.0	406	DHE4_PSVT1	O00128 ictalurid h
26	26	100.0	447	DHE4_PSVT1	O951f9 psychobact
27	26	100.0	471	STCH_HUMAN	P46723 homo sapien
28	26	100.0	501	FLAA_AOUPY	P46723 aquilifex pyr
29	26	100.0	504	BCHB_ACTRU	O954b5 acidiphiliu
30	26	100.0	557	YOHK_HUCSU	P54509 bacillus su
31	26	100.0	1084	HDA4_HUMAN	P54524 homo sapien
32	26	100.0	1178	RPOB_TREPA	O83269 treponema p
33	26	100.0	1275	COBN_PSEDE	P29929 pseudomonas

34	24	92.3	1409	1	HAP1_HAETN	P44596 haemophilus
35	23	88.5	26	1	AMD1_CHICK	P81073 gallus gall
36	23	88.5	108	1	KVSD_MOUSE	P01636 mus musculu
37	23	88.5	185	1	ADBP_BACSU	P54570 bacillus su
38	23	88.5	191	1	Y777_METVA	O58187 methanococc
39	23	88.5	295	1	DAFP_METVA	O58189 methanococc
40	23	88.5	299	1	Y779_METVA	P75794 escherichia
41	23	88.5	308	1	PFLE_ECOS7	O00119 ictalurid h
42	23	88.5	308	1	VG53_HSV1	P04194 clostridium
43	23	88.5	319	1	DCHS_CLOPE	O57809 pyrococcus
44	23	88.5	325	1	LAID_PYROH	O67189 aquilifex aeo
45	23	88.5	350	1	YB04_AOUAE	

ALIGNMENTS

RESULT 1	Y44A_MYCPN	STANDARD:	PRT: 277 AA.
ID	Y44A_MYCPN		
AC	P75151:		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical lipoprotein MG440 homolog 1 precursor (E09_orf277).		
GN	MPN646 OR MP196.		
OS	Mycoplasma pneumoniae.		
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.		
OX	NCBI_TaxID=2104;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 29342 / M129;		
RX	MEDLINE=97105885; PubMed=6948633;		
RA	Himmelfreih R., Hillbert H., Plagens H., Pirkl E., Li B.-C.,		
RA	Herrmann R.;		
RT	"Complete sequence analysis of the genome of the bacterium Mycoplasma		
RT	pneumoniae.";		
RL	Nucleic Acids Res. 24:4420-4449(1996).		
CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor		
CC	(potential).		
CC	-1- SIMILARITY: BELONGS TO THE MG439 / MG440 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL: AE000020: AAB95844.1; -		
DR	InterPro: IPR001585; Lipoprotein_3.		
DR	Pfam: PF00938; Lipoprotein_3; 1.		
DR	ProDom: PD003276; Lipoprotein_3; 1.		
DR	PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.		
KW	Hypothetical protein; Lipoprotein; Membrane; Signal;		
KW	Complete proteome.		
FT	SIGNAL	1	25
FT	CHAIN	26	277
FT	LIPID	26	26
SO	SEQUENCE	277 AA; 31097 MW; A17DC45BED6D3065 CRC64;	
Query Match	100.0%;	Score 26;	DB 1; Length 277;
Best Local Similarity	100.0%;	Pred. No. 39;	
Matches	5;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 LYDAK 5		
Db			
	85 LYDAK 89		
RESULT 2	SPB7_HUMAN		

```

ID  SEP7_HUMAN      STANDARD:      PRT:      380 AA.
AC  075635;
DT  15-JUN-2002 (Rel. 41, Created)
DR  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Megsin (TFP55) (Serpin B7).
GN  SERPINB7.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=97326116; PubMed=9182567;
RA  Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
RA  Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA  Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA  Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,
RT  "Purification, cDNA cloning, and characterization of a new serpin with
RT  megakaryocyte maturation activity."
RT  J. Biol. Chem. 272:15373-15380(1997).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  TISSUE=Mesangial cells;
RA  MEDLINE=98376492; PubMed=9710452;
RA  Miyata T., Nangaku K., Suzuki D., Inagi R., Uregami K., Sakai H.,
RA  Okubo K., Kurokawa K.;
RT  "A mesangium-predominant gene, megin, is a new serpin upregulated in
RT  IgA nephropathy."
RT  J. Clin. Invest. 102:828-836(1998).
CC  -1- FUNCTION: Might function as an inhibitor of lys-specific
CC  proteases. Might influence the maturation of megakaryocytes via
CC  its action as a serpin.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC  -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@isb-sdb.ch).
CC  -----
DR  EMBL; D88575; BAA31232.1; -
DR  EMBL; AF027866; AAC64506.1; -
DR  HSSP; P05619; 1HE.
DR  Genew; HGNC:13902; SERPINB7.
DR  MIM; 603357; -
DR  InterPro; IPR000215; Serpin.
DR  Pfam; PF00079; serpin; 1.
DR  SMART; SM00093; SERPIN; 1.
DR  PROSITE; PS00284; SERPIN; 1.
KW  Serpin; Serine protease inhibitor.
FT  ACT_SITE 347 348 REACTIVE BOND (BY SIMILARITY).
SQ  SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CF605 CRC64;

```

```

Query Match      100.0%; Score 26; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  1 LYDAK 5
    |||||
DB  121 LYDAK 125

```

```

RESULT 3
ID  YJ95_YEAST      STANDARD:      PRT:      408 AA.
AC  P47160;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)

```

```

DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  Hypothetical 45.1 kDa protein in RP55-ZMS1 intergenic region.
GN  YJRI25C OR J2048.
OS  Saccharomyces cerevisiae (Baker's Yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Rose M., Koetter P., Entian K.D.;
RT  Submitted (SEP-1995) to the EMBL/Genbank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=97326116; PubMed=9182567;
RA  Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
RA  Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA  Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA  Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,
RT  "Purification, cDNA cloning, and characterization of a new serpin with
RT  megakaryocyte maturation activity."
RT  J. Biol. Chem. 272:15373-15380(1997).
RN  [3]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=98376492; PubMed=9710452;
RA  Miyata T., Nangaku K., Suzuki D., Inagi R., Uregami K., Sakai H.,
RA  Okubo K., Kurokawa K.;
RT  "A mesangium-predominant gene, megin, is a new serpin upregulated in
RT  IgA nephropathy."
RT  J. Clin. Invest. 102:828-836(1998).
CC  -1- FUNCTION: Might function as an inhibitor of lys-specific
CC  proteases. Might influence the maturation of megakaryocytes via
CC  its action as a serpin.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC  -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@isb-sdb.ch).
CC  -----
DR  EMBL; 249625; CAA89656.1; -
DR  SGD; S0003886; YJRI25C.
DR  InterPro; IPR001026; ENTH.
DR  Pfam; PF01417; ENTH; 1.
DR  SMART; SM00273; ENTH; 1.
KW  Hypothetical protein.
SQ  SEQUENCE 408 AA; 45091 MW; D666ECB1C0D074FE CRC64;

```

```

Query Match      100.0%; Score 26; DB 1; Length 408;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  1 LYDAK 5
    |||||
DB  12 LYDAK 16

```

```

RESULT 4
ID  MCRB_METVA      STANDARD:      PRT:      443 AA.
AC  P07856;
DT  01-AUG-1988 (Rel. 08, Created)
DT  01-AUG-1988 (Rel. 08, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Methyl-coenzyme M reductase beta subunit (EC 1.8.-.-).
GN  MCRB.
OS  Methanococcus vannielii.
OC  Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC  Methanococcaceae; Methanococcus.
OX  NCBI_TaxID=2187;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Cram D.S., Sherf B.A., Libby R.T., Mattaliano R.J., Ramachandran K.L.,
RA  Reeve J.N.;
RT  "Structure and expression of the genes, mcrBDCa, which encode the
RT  subunits of component C of methyl coenzyme M reductase in
RT  Methanococcus vannielii."
RT  Proc. Natl. Acad. Sci. U.S.A. 84:3992-3996(1987).
CC  -1- FUNCTION: Reduction of methyl-coenzyme M (2-(methylthio)
CC  ethanesulfonic acid) with 7-mercaptoheptanoylthreonine phosphate
CC  to methane and an heterodisulfide.
CC  -1- CATALYTIC ACTIVITY: CH(3)-S-CoM + H-S-HTP = CH(4) + CoM-S-S-HTP.
CC  -1- COFACTOR: THE ENZYME COMPLEX BINDS TIGHTLY (BUT NOT COVALENTLY)
CC  TO TWO MOLECULES OF COENZYME F430. F430 IS A YELLOW NICKEL
CC  PORPHINOID.
CC  -1- PATHWAY: Methanogenesis; last step.
CC  -1- SUBUNIT: HEXAMER OF TWO ALPHA, TWO BETA, AND TWO GAMMA CHAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M16893; AAA72594.1; -.

DR PIR: A27793; A27793.

DR HSSP: P11560; 1MRO.

DR InterPro: IPR003179; MCR_beta.

DR Pfam: PF02241; MCR_beta.1.

DR Pfam: PF02783; MCR_beta.N.1.

KW Methanogenesis; Oxidoreductase.

SQ SEQUENCE 443 AA; 46954 MW; 7666A0D427A98517 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 10 LYDAK 14

RESULT 5
TOLC_SALEN
ID TOLC_SALEN STANDARD; PRT; 491 AA.
AC 054001;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Outer membrane protein tolC precursor.
GN TOLC.
OS Salmonella enteritidis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD5;
RX MEDLINE=96111490; PubMed=8801424;
RA Stone B.J., Miller V.L.;
RT Salmonella enteritidis has a homologue of tolC that is required for virulence in BALB/c mice."
RL Microbiol. 17:701-712(1995).
CC -1- FUNCTION: REQUIRED FOR PROPER EXPRESSION OF OUTER MEMBRANE PROTEIN GENES SUCH AS OMP, NMPC, PROTEIN 2, HEMOLYSIN, COLICIN V, OR COLICIN E1. MAY BE SPECIALIZED FOR SIGNAL SEQUENCE INDEPENDENT, EXTRACELLULAR SECRETION IN GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: Outer membrane.
CC -1- SIMILARITY: BELONGS TO THE PRP FAMILY OF SECRETION PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U25178; AAC43973.1; -.

DR HSSP: P02930; 1EK9.

DR InterPro: IPR003423; OEP.

DR Pfam: PF02321; OEP.2.

KW Transport; Outer membrane; Signal.

FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 491 OUTER MEMBRANE PROTEIN TOLC.
SQ SEQUENCE 491 AA; 53725 MW; 58F0FB4C0257FA4 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 400 LYDAK 404

RESULT 6
BCHB_RHOGE
ID BCHB_RHOGE STANDARD; PRT; 543 AA.
AC 09UPA3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Light-independent protochlorophyllide reductase subunit B (EC 1.18.-.-) (Li-POR subunit B).
GN BCHB.
OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Rubrivivax.
OX NCBI_TaxID=28068;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL144;
RX MEDLINE=21240759; PubMed=11343129;
RA Igarashi N., Harada J., Nagashima S., Matsuura K., Shinada K., Nagashima K.V.;
RT "Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria."
RL J. Mol. Evol. 52:333-341(2001).
CC -1- FUNCTION: Uses Mg-ATP and reduced ferredoxin to reduce ring D of protochlorophyllide (pchlide) to form chlorophyllide a (Chlide) (By similarity). This reaction is light-independent.
CC -1- PATHWAY: Light-independent bacteriochlorophyll biosynthesis.
CC -1- SUBUNIT: Protochlorophyllide reductase is thought to be composed of three subunits: bchl, bchN and bchB. could form a heterotrimer of two bchl and two bchN subunits.
CC -1- SIMILARITY: BELONGS TO THE CHL / BCHB / BCHZ FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AB034704; BAA94058.1; -.

DR InterPro: IPR000510; Oxired_nitrogensel.

DR Pfam: PF00148; Oxidored_nitro; 1.

KW TIGRPFAMs: TIGR01278; DPOR_BchB; 1.
KW Oxidoreductase; Photosynthesis; Bacteriochlorophyll biosynthesis.

SQ SEQUENCE 543 AA; 58697 MW; 76ADF68BA2C1D067 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 543;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
|||||
DB 534 LYDAK 538

RESULT 7
ASNI_YEAST
ID ASNI_YEAST STANDARD; PRT; 571 AA.
AC P49089;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Asparagine synthetase [glutamine-hydrolyzing] 1 (EC 6.3.5.4)
DE (Glutamine-dependent asparagine synthetase 1).
GN ASNI OR YPR145W OR P4659.3.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]

```

RP SEQUENCE FROM N.A.
RA MEDLINE-97109535; PubMed-8951815;
RA Dang V.D., Valens M., Bolotin-Fukuhara M., Daignan-Formier B.;
RT "Cloning of the ASN1 and ASN2 genes encoding asparagine synthetases
RT in Saccharomyces cerevisiae: differential regulation by the
RT CCAT-box-binding factor.";
RL Mol. Microbiol. 22:681-692(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-5288C / AB972;
RX MEDLINE-97131371; PubMed-9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
RA Arayo R., Aparicio A., Barrall B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., Dipaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hildert H., Hillier L.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kahan S., Kleine K.,
RA Komp C., Kurd O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Mueller-Auer S., Namath A., Newtlich U., Oeffner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schaefer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettein H.,
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambolt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hanl J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + L-glutamine = AMP +
CC -1- phosphate + L-asparagine + L-glutamate.
CC -1- PATHWAY: Asparagine biosynthesis.
CC -1- SIMILARITY: CONTAINS 1 TYPE-2 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC -1- SIMILARITY: TO OTHER ASN SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC in the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48675; CAAB8594.1; -
DR EMBL: U40829; AAB68284.1; -
DR HSSP: P22106; ICT9.
DR SGD: S0006349; ASN1.
DR InterPro: IPR001962; Asn_synthase.
DR InterPro: IPR000583; GATase_2.
DR Pfam: PF00103; GATase_2; 1.
DR Pfam: PF00733; Asn_synthase; 1.
DR PROSITE: PS00443; GATASE_TYPE_II; 1.
KW Ligase: Asparagine biosynthesis; Glutamine amidotransferase;
KW Multigene family.
FT INT_MET 0
FT ACT_SITE 1 BY SIMILARITY.
FT GATASE (BY SIMILARITY).
SQ SEQUENCE 571 AA; 64339 MW; 1996772387502E40 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 571;
Best Local Similarity 100.0%; Pred. NO. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LYDAR 5
DB 123 LYDAR 127

```

```

DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
GN E1.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10587;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94265501; PubMed-8205838;
RA Dellus H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-92124753; PubMed-1310198;
RA Marich J.E., Ponteler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
RT human papillomavirus type 35.";
RL Virology 186:770-776(1992).
CC -1- FUNCTION: ATP-DEPENDENT. IT FORMS A COMPLEX WITH THE VIRAL E2
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC in the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74477; CAAS2563.1; -
DR EMBL: M74117; AAA56968.1; -
DR PIR: A40824; WIML35.
DR PIR: S36523; S36523.
DR InterPro: IPR001177; Papillom_E1.
DR Pfam: PF00519; E1; 1.
DR Pfam: PF00524; E1_N; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
FT NP_BIND 463 470
FT DEVS -> SSV (IN REF. 2).
FT CONFICT 31 34
FT CONFICT 142 143
FT CONFICT 235 258
FT CONFICT 269 281
FT CONFICT 417 418
FT CONFICT 515 552
FT CONFICT 567 587
FT CONFICT 613 613
FT CONFICT 72122 MW; 3C694D4451791003 CRC64;
SQ SEQUENCE 637 AA; 72122 MW; 3C694D4451791003 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 637;
Best Local Similarity 100.0%; Pred. NO. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LYDAR 5
DB 499 LYDAR 503

```

```

RESULT 8
ID VE1_HPV35 STANDARD: PRT: 637 AA.
AC P27220;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

RESULT 9
ID NODO_RHIME STANDARD: PRT: 641 AA.
AC P13442;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)

```

15-JUN-2002 (Rel. 41, last annotation update)
 DE NodQ bifunctional enzyme (Nodulation protein O) [includes: Sulfate
 DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylylate
 DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
 DE kinase (EC 2.7.1.35) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
 DE phosphotransferase)].
 GN NODQ OR RA0469 OR SMA0857.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 ON NCBI_TaxID=382;
 RP SEQUENCE FROM N.A.
 RC STRAIN=RCR2011 / SU47;
 RX MEDLINE=89313304; PubMed=2546009;
 RA Cervantes E., Sharma S.B., Maillet F., Vasse J., Truchet G.,
 RA Rosenberg C.;
 RT "The Rhizobium meliloti host range nodQ gene encodes a protein which
 RT shares homology with translation elongation and initiation factors.";
 RL Mol. Microbiol. 3:745-755(1989).
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92288403; PubMed=2520820;
 RA Schwedock J., Long S.R.;
 RT "Nucleotide sequence and protein products of two new nodulation genes
 RT of Rhizobium meliloti, nodP and nodQ.";
 RL Mol. Plant Microbe Interact. 2:181-194(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barby-Hubler F., Bowser L., Capela D., Gallbert F., Guzy J.,
 RA Gurjal M., Hong A., Huizer L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Kohn K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSyma megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 CC -1- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
 CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
 CC SULFURYLASE ACTIVITY (BY SIMILARITY).
 CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + sulfate -> diphosphate + adenylylsulfate.
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate -> ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -1- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODO, MAY BE
 CC PHYSICALLY ASSOCIATED (POTENTIAL).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE GTP-BINDING
 CC ELONGATION FACTOR FAMILY. CTSN/NODO SUBFAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE APS KINASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X14809; CAA32914.1; -
 DR EMBL: M68858; AAA26343.1; -
 DR EMBL: AEO07237; AAK65127.1; -
 DR PIR: S14899; Z2ZRNQ.
 DR InterPro: IPR002891; APS_kinase.
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR000795; EFTU_D2.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR Pfam: PF01583; APS_kinase; 1.

Pfam: PF03144; GTP_EFTU_D2; 1.
 DR ProDom: PD002350; APS_kinase; 1.
 DR TIGRfams: TIGR00231; small_gtp; 1.
 DR TIGRfams: TIGR00455; apsk; 1.
 DR PROSITE: PS00301; EFATOR_GTP; 1.
 KW Phosite: Nodulation; Transferase; Nucleotidyltransferase; GTP-binding;
 KM Kinase; ATP-binding; Multifunctional enzyme; Complete proteome.
 FT DOMAIN 1 458
 FT DOMAIN 459 641
 FT NE_BIND 31 38
 FT NE_BIND 110 114
 FT NE_BIND 165 168
 FT NE_BIND 467 474
 FT ACT_SITE 524 524
 FT ACT_SITE 524 524
 FT SEQUENCE 641 AA; 70614 MW; 1E1261F04ED33A93 CRC64;
 SQ
 Query Match 100.0%; Score 26; DB 1; Length 641;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LVDAA 5
 Db 45 LVDAA 49
 RESULT 10
 NCAL_CHICK STANDARD; PRT; 1091 AA.
 ID NCAL_CHICK
 AC P13590; Q90919; Q90918;
 DT 01-JUN-1998 (Rel. 13, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM
 DE 180).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE OF 1-175 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=87206190; PubMed=3576199;
 RA Cunningham B.A., Hemperly J.J., Murray B.A., Prediger E.A.,
 RA Brackenbury R., Edelman G.M.;
 RT "Neural cell adhesion molecule: structure, immunoglobulin-like
 RT domains, cell surface modulation, and alternative RNA splicing.";
 RL Science 236:799-806(1987).
 RN [2]
 RP SEQUENCE OF 128-1091 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=86206089; PubMed=3458261;
 RA Hemperly J.J., Murray B.A., Edelman G.M., Cunningham B.A.;
 RT "Sequence of a cDNA clone encoding the polysialic acid-rich and
 RT cytoplasmic domains of the neural cell adhesion molecule N-CAM.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3037-3041(1986).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOPFORM D).
 RX MEDLINE=87092340; PubMed=3467341;
 RA Hemperly J.J., Edelman G.M., Cunningham B.A.;
 RT "cDNA clones of the neural cell adhesion molecule (N-CAM) lacking a
 RT membrane-spanning region consistent with evidence for membrane
 RT attachment via a phosphatidylinositol intermediate.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:9822-9826(1986).
 RN [4]
 RP SEQUENCE OF 810-1069 FROM N.A.
 RX MEDLINE=87033934; PubMed=3771645;
 RA Murray B.A., Owens G.C., Prediger E.A., Crossin K.L.,
 RA Cunningham B.A., Edelman G.M.;
 RT "Cell surface modulation of the neural cell adhesion molecule
 RT resulting from alternative mRNA splicing in a tissue-specific
 RT developmental sequence.";
 RL J. Cell Biol. 103:1431-1439(1986).
 RN [5]


```

OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=M303;
RX MEDLINE=94186543; PubMed=8138573;
RA WOZNIAK R.W., BLOBEL G., ROUT M.P.;
RT POM152 is an integral protein of the pore membrane domain of the
RT yeast nuclear envelope.
RL Cell Biol. 125:31-42(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA BADCOCK K., CHURCHER C., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE NUCLEAR PORE COMPLEX WHICH MAY PROVIDE
CC THE NECESSARY ASYMMETRY REQUIRED FOR ANCHORING STRUCTURES SUCH AS
CC CYTOPLASMIC FILAMENTS AND THE NUCLEOPLASMIC CAGE.
CC -1- SUBUNIT: INTERACTS WITH NUP170 AND/OR NUP188.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. NUCLEAR PORE
CC MEMBRANE.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: TO S.POMBE SPEC29A10.07.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z31592; CAA83459.1; -
DR EMBL; Z48622; CAA88554.1; -
DR PIR; A53824; A53824.
DR SGD; S0004736; POM152.
KM Nuclear protein; Transmembrane; Repeat; Glycoprotein.
FT DOMAIN 1 175 PORE SIDE (POTENTIAL).
FT TRANSMEM 176 195 POTENTIAL.
FT DOMAIN 196 1337 CISTERNAL SIDE (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC...);
FT DOMAIN 390 1276 8 X 24 AA APPROXIMATE REPEATS.
FT REPEAT 390 413 1.
FT REPEAT 626 650 2.
FT REPEAT 732 755 3.
FT REPEAT 836 859 4.
FT REPEAT 943 966 5.
FT REPEAT 1058 1077 6.
FT REPEAT 1157 1178 7.
FT REPEAT 1253 1276 8.
SQ SEQUENCE 1337 AA; 151651 MW; A024A42069193898 CRC64;

Query Match 100.0%; Score 26; DB 1; Length 1337;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., MATANABE A., YAMADA M., YASUDA M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- SIMILARITY: BELONGS TO THE CAPAB / TERDEZ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90903; BAA17063.1; -
DR InterPro; IPR003325; Terd.
DR Pfam; PF02342; Terd; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 179 AA; 19422 MW; 19BC13A5C710FEA7 CRC64;

Query Match 92.3%; Score 24; DB 1; Length 179;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 LYDAK 5
Db 662 LYDAK 666

RESULT 12
ID YH64_SYNY3 STANDARD; PRT; 179 AA.
AC P73042;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein sir1764.
GN SLR1764.
OS Synechocystis sp. (strain PCC 6803).

```

```

OY 1 LYDAK 5
Db 100 LYDAK 104

RESULT 13
ID Y418_METJA STANDARD; PRT; 257 AA.
AC O57861;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0418.
GN MJ0418.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA BILT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.T.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERICK J.M., GLODEK A.,
RA SCOTT J.L., GEORGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UUTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KALNE B.P., BORDOVSKY M.,
RA KIENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U67494; AAB98416.1; -
CC TIGR: M00418; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 7 POTENTIAL.
SQ SEQUENCE 257 AA; 28765 MW; CFEAIAA6AB38000 CRC64;

Query Match
Best Local Similarity 92.3%; Score 24; DB 1; Length 257;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 82 IYDAK 86

RESULT 14

ID XDBH_ECO57 STANDARD; PRT; 292 AA.
AC 08X6C5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Xanthine dehydrogenase, FAD binding subunit (EC 1.1.1.204).
GN XDBH OR 24206 OR ECS3740.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Link S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).

RT Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tode T.,
RA Ikeda T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuwara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).

CC -1- FUNCTION: Presumed to be a dehydrogenase, but possibly an oxidase.
CC Participates in limited purine salvage (requires aspartate) but
CC does not support aerobic growth on purines as the sole carbon
CC source (purine catabolism) (By similarity).
CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
CC -1- COFACTOR: FAD (By similarity).
CC -1- PATHWAY: Purine catabolism; first committed step.
CC -1- SUBUNIT: Heterotrimer of xdbh, xdbb and xdcn (Probable).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: AE005516; AAG57996.1; -
DR EMBL: AP002563; BAB37163.1; -

KW Oxidoreductase; NAD; Flavoprotein; FAD; Purine metabolism;
KW Purine salvage; Complete proteome.
SQ SEQUENCE 292 AA; 31561 MW; FEC44F990BF9BC1 CRC64;

Query Match
Best Local Similarity 92.3%; Score 24; DB 1; Length 292;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 130 IYDAK 134

RESULT 15

ID XDBH_ECOLI STANDARD; PRT; 292 AA.
AC 046800;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Xanthine dehydrogenase, FAD binding subunit (EC 1.1.1.204).
GN XDBH OR B2867.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).

RT Science 277:1453-1474(1997).
RN [2]
RP DISCUSSION OF FUNCTION.
RC STRAIN-K12 / W3110;
RX MEDLINE=20444178; PubMed=10986234;
RA Xi H., Schneider B.L., Reitzer L.;
RT "Purine catabolism in Escherichia coli and function of xanthine
RT dehydrogenase in purine salvage."
RL J. Bacteriol. 182:5332-5341(2000).

CC -1- FUNCTION: Presumed to be a dehydrogenase, but possibly an oxidase.
CC Participates in limited purine salvage (requires aspartate) but
CC does not support aerobic growth on purines as the sole carbon
CC source (purine catabolism).
CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.
CC -1- COFACTOR: FAD (By similarity).
CC -1- PATHWAY: Purine catabolism; first committed step.
CC -1- SUBUNIT: Heterotrimer of xdbh, xdbb and xdcn (Probable).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U28375; AAB83048.1; -
DR EMBL: AE000370; AAC75905.1; -
DR Ecogene; EGI3050; xdbb.
DR InterPro: IPR005107; CO_deh_flav_C.
DR InterPro: IPR002346; dehydrog_molyb.
DR Pfam: PF00941; FAD_binding_5; 1.
DR Pfam: PF03450; CO_deh_flav_C; 1.
KW Oxidoreductase; NAD; Flavoprotein; FAD; Purine metabolism;
KW Purine salvage; Complete proteome.
SQ SEQUENCE 292 AA; 31556 MW; 74A07D137DA857E8 CRC64;

Query Match
92.3%; Score 24; DB 1; Length 292;

Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
:||||
Db 130 TYDAK 134

Search completed: July 11, 2003, 11:55:18
Job time : 1.47817 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 2.18295 Seconds
(without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-7
Perfect score: 26
Sequence: 1 LYDAK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTRMBL_21: *
1: sp.archaea: *
2: sp.bacteria: *
3: sp.fungi: *
4: sp.human: *
5: sp.invertebrate: *
6: sp.mammal: *
7: sp.mhc: *
8: sp.organelle: *
9: sp.phage: *
10: sp.plant: *
11: sp rodent: *
12: sp.virus: *
13: sp.vertebrate: *
14: sp.unclassified: *
15: sp.virus: *
16: sp.bacteriaph: *
17: sp.archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	100.0	107	17	0973B2
2	26	100.0	108	9	0971J3
3	26	100.0	135	5	08SUH3
4	26	100.0	175	5	001760
5	26	100.0	178	16	066796
6	26	100.0	195	17	08U3H5
7	26	100.0	197	17	08B353
8	26	100.0	199	5	022438
9	26	100.0	213	10	09ASL1
10	26	100.0	230	12	088960
11	26	100.0	250	2	09PCV9
12	26	100.0	250	5	08T3H1
13	26	100.0	253	16	0926W7
14	26	100.0	259	16	09PJ15
15	26	100.0	281	3	09P7L6
16	26	100.0	288	10	09CA01

17	26	100.0	299	17	08TKS0	08tkso methanarac
18	26	100.0	301	16	097GN2	097gn2 clostridium
19	26	100.0	304	17	097V63	097v63 sulfolobus
20	26	100.0	306	2	0936E9	0936e9 staphylococ
21	26	100.0	329	2	0936V9	0936v9 selenomonas
22	26	100.0	331	9	094M05	094m05 bacterioph
23	26	100.0	334	2	08RPD7	08rpd7 legionella
24	26	100.0	341	16	091072	091072 pseudomonas
25	26	100.0	354	12	096FM3	096fm3 turkey herp
26	26	100.0	354	12	067629	067629 marek s dis
27	26	100.0	356	16	09RRJ2	09rrj2 deinococcus
28	26	100.0	357	10	09FJ41	09fj41 arabidopsis
29	26	100.0	370	16	08XK23	08xk23 clostridium
30	26	100.0	377	16	098509	098509 rhizobium l
31	26	100.0	388	2	09KX77	09kx77 staphylococ
32	26	100.0	401	13	090976	090976 gallus gall
33	26	100.0	409	16	098IS9	098is9 rhizobium l
34	26	100.0	421	16	09EV35	09ev35 staphylococ
35	26	100.0	428	16	098M33	098m33 rhizobium l
36	26	100.0	430	16	09R1V4	09r1v4 streptomyce
37	26	100.0	439	16	08U6J7	08u6j7 agrobacteri
38	26	100.0	448	17	0974Y6	0974y6 sulfolobus
39	26	100.0	450	13	090809	090809 gallus gall
40	26	100.0	460	5	08WS95	08ws95 glossina mo
41	26	100.0	464	10	094ETO	094eto hevea brasl
42	26	100.0	464	10	0944F8	0944f8 hevea brasl
43	26	100.0	474	10	093773	093773 pinus sylve
44	26	100.0	475	5	094404	094404 caenorhabd
45	26	100.0	491	16	082LZ4	082l24 salmone

ALIGNMENTS

RESULT 1
ID: 0973B2 PRELIMINARY; PRT: 107 AA.
AC: 0973B2;
DT: 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DE: 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DT: Putative transcriptional regulator.
GN: SP0980.
OS: Sulfolobus tokodaii.
OC: Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC: Sulfolobus.
OX: NCBI_TaxID=111955;
RN: [1]
RP: SEQUENCE FROM N.A.
RC: STRAIN=JCM 10545 / 7;
RX: Pubmed=11572479;
RA: Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya N., Kato Y., Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kusuda N., Oguchi A., Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A., Oshima T., Kikuchi H.;
RA: Complete genome sequence of an aerobic thermophilic
RT: Crenarchaeon, Sulfolobus tokodaii strain7.";
RL: DNA Res. 8:123-140(2001).
DR: EMBL: AP000984; BAB6001.1;
KW: Hypothetical protein; Complete proteome.
SQ: SEQUENCE 107 AA; 12386 MW; 2CF24E31B02D3C68 CRC64;

Query Match 100.0%; Score 26; DB 17; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 5 LYDAK 9

```

RESULT 2
Q971J3      PRELIMINARY;      PRT;      108 AA.
ID 0971J3
AC 0971J3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Rad protein.
GN RAD.
OS Lactobacillus bacteriophage phi adh.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
ON NCBI_TaxID:12417;
RN [1]
RP SEQUENCE FROM N.A.
RA Altermann E.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases..
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-9384014; PubMed-10452953;
RA Altermann E., Klein J., Henrich B.;
RT "Primary structure and features of the genome of the Lactobacillus gasseri temperate bacteriophage phi-adh.";
RL Gene 236:333-346(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-95138034; PubMed-7836307;
RA Henrich B., Binschofer B., Blesl U.;
RT "Primary structure and functional analysis of the lysis genes of Lactobacillus gasseri bacteriophage phi-adh.";
RL J. Bacteriol. 177:723-732(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE-93231538; PubMed-8472961;
RA Fremux C., De Antoni G., Raya R., Kleenhammer T.;
RT "Genetic organization and sequence of the region encoding integrative functions from Lactobacillus gasseri temperate bacteriophage phi-adh.";
RL Gene 126:61-66(1993).
RN [5]
RP SEQUENCE FROM N.A.
RA Engel G., Altermann E., Klein J., Henrich B.;
RT "Structure of a genome region of the Lactobacillus gasseri temperate phage phi adh covering a repressor gene and cognate promoters.";
RL Gene 210:67-70(1998).
DR EMBL: AJ131519; CAB52483.1;
DR InterPro: IPR001387; HTH_3.
DR Pfam: PF01381; HTH_3.1.
DR SMART: SM00530; HTH_XRE; 1.
SQ SEQUENCE 108 AA; 12086 MW; EC9109B3318BC5E2 CRC64;

Query Match      100.0%; Score 26; DB 9; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 104 LYDAK 108

RESULT 3
Q85UH3      PRELIMINARY;      PRT;      135 AA.
ID 085UH3;
AC 085UH3;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical protein EC010_0230.
GN EC010_0230.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
ON NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21576510; PubMed-11719806;
RA Kalinka M.D., Duprat S., Cornilliot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyreallade E., Brottier P., Mincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivas C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL: AL590449; CAD25742.1;
KM Hypothetical protein.
SQ SEQUENCE 135 AA; 16153 MW; A202A9AD801B1B5F CRC64;

Query Match      100.0%; Score 26; DB 5; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 87 LYDAK 91

RESULT 4
O01760      PRELIMINARY;      PRT;      175 AA.
ID 001760;
AC 001760;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 19.3 kDa protein.
GN F55A12.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA "The sequence of C. elegans cosmid F55A12.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003130; AAB54128.2;
KM Hypothetical protein.
SQ SEQUENCE 175 AA; 19251 MW; 443B949A9EEF2897B CRC64;

Query Match      100.0%; Score 26; DB 5; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 54 LYDAK 58

```

RESULT 5
 066796 PRELIMINARY; PRT; 178 AA.
 ID 066796
 AC 066796;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein AQ_507.
 GN AQ_507.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 NC NCB1_TaxID=63363;
 OX NCB1_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL: AE000693; AAC06761.1; -;
 DR InterPro: IPR002934; NTP_transf.
 DR InterPro: IPR001201; PAP_25A_core.
 DR Pfam: PF01909; NTP_transf.2; 1.
 KM Hypothetical protein: Complete proteome.
 SQ SEQUENCE 178 AA; 20421 MW; C3EF9688B9A78F42 CRC64;

Query Match 100.0%; Score 26; DB 16; Length 178;
 Best Local Similarity 100.0%; Pred. NO. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 |||||
 Db 61 LYDAK 65

RESULT 6
 0803H5 PRELIMINARY; PRT; 195 AA.
 ID 0803H5
 AC 0803H5;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Transcription initiation factor TFIIIE.
 GN PF0491.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 NC NCB1_TaxID=2261;
 OX NCB1_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEF-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE010171; AAL80615.1; -;
 KW Initiation factor: Complete proteome.
 SQ SEQUENCE 195 AA; 22885 MW; F7507AF9F3353763 CRC64;

Query Match 100.0%; Score 26; DB 17; Length 195;
 Best Local Similarity 100.0%; Pred. NO. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 |||||
 Db 58 LYDAK 62

RESULT 7
 058353 PRELIMINARY; PRT; 197 AA.
 ID 058353

AC 058353;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical protein PH0619.
 GN PH0619.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 NC NCB1_TaxID=53953;
 OX NCB1_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000003; BAA29708.1; -;
 DR InterPro: IPR002853; TFIIIE_alpha.
 DR InterPro: IPR005241; TFIIIE_alpha-rel.
 DR Pfam: PF02002; TFIIIE_alpha; 1.
 DR SMART: SM00531; TFIIIE; 1.
 DR TIGRfams: TIGR00373; TIGR00373; 1.
 KM Hypothetical protein: Complete proteome.
 SQ SEQUENCE 197 AA; 23127 MW; 2A0AA7110389AF61 CRC64;

Query Match 100.0%; Score 26; DB 17; Length 197;
 Best Local Similarity 100.0%; Pred. NO. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 |||||
 Db 62 LYDAK 66

RESULT 8
 022438 PRELIMINARY; PRT; 199 AA.
 ID 022438
 AC 022438;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE T12G3.5 protein.
 GN T12G3.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 OX NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sims M.A.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z68752; CAA92984.1; -;
 SQ SEQUENCE 199 AA; 23636 MW; A4FF219673154006 CRC64;

Query Match 100.0%; Score 26; DB 5; Length 199;
 Best Local Similarity 100.0%; Pred. NO. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5

DB 134 LYDAK 138

RESULT 9

09ASL1 PRELIMINARY; PRT; 213 AA.
ID 09ASL1;
AC 09ASL1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE P0439B06.1 protein (P0482C06.29 protein).
GN P0439B06.1 OR P0482C06.29.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;

RM (1)
RN SEQUENCE FROM N.A.
RC STRAIN-CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone: P0439B06.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RM (2)
RN SEQUENCE FROM N.A.
RC STRAIN-CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone: P0482C06.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002882; BAB39866.1;
DR EMBL; AP002845; BAB76632.1;
DR InterPro: IPR000169; SHPOrt.acsite.
DR PROSITE: PS00639; THTOL_PROTEASE_HIS; UNKNOWN_1.
SQ SEQUENCE 213 AA; 22915 MW; 9EAA4FED9184AA3 CRC64;

Query Match 100.0%; Score 26; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 67 LYDAK 71

RESULT 10

088960 PRELIMINARY; PRT; 230 AA.

AC 088960;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE DNA polymerase (Fragment).
OS Unidentified phycodnavirus clone OT05.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae;
OC Environmental samples.
OX NCBI_TaxID=43245;

RM (1)
RN SEQUENCE FROM N.A.
RC MEDLINE=96316404; Pubmed=8702280;
RA Chen F., Suttle C.A., Short S.M.;
RT "Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes";
RL Appl. Environ. Microbiol. 62:2869-2874(1996).
-1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -> N DIPHOSPHATE + (DNA)(N).

CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC EMBL; U36935; AAB18299.1;
CC InterPro: IPR002064; DNA_pol_B.
DR Pfam: PF00136; DNA_pol_B; 2.
DR PRINTS; PR00106; DNAPOB.

DR PROSITE: PS00116; DNA_POLYMERASE_B; UNKNOWN_1.
KW DNA replication; DNA-binding; DNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 230
SQ SEQUENCE 230 AA; 25543 MW; 983319E04EC18388 CRC64;

Query Match 100.0%; Score 26; DB 12; Length 230;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 104 LYDAK 108

RESULT 11

09FCV9 PRELIMINARY; PRT; 250 AA.

AC 09FCV9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Para protein.
GN PARA.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;

RM (1)
RN SEQUENCE FROM N.A.
RC STRAIN-EGD;
RA Brehm K., Foerster S., Kretz J.;
RT "The L. monocytogenes parv locus.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ010494; CAC06097.1;
DR InterPro: IPR000707; ATPase_PARA.
DR InterPro: IPR000392; NitrogenaseII.
DR Pfam: PF00991; Para; 1.
DR PRINTS; PR00091; NITROGNASEII.
SQ SEQUENCE 250 AA; 27186 MW; A77DA30B28C94CDD CRC64;

Query Match 100.0%; Score 26; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
DB 228 LYDAK 232

RESULT 12

08T3H1 PRELIMINARY; PRT; 250 AA.

AC 08T3H1;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hypothetical 27.7 kDa protein.
GN F55A12.2.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;

RM (1)
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MEDLINE=99069613; Pubmed=9851916;
RA Waterston R.;

RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RM (2)
RN SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
 RA Pauley A., Gattung S.;
 RT "The sequence of *C. elegans* cosmid F55A12.";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN 131
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF001310; AM15579.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 250 AA; 27678 MW; BCD8E30A54E52215 CRC64;

Query Match 100.0%; Score 26; DB 5; Length 250;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 Db 54 LYDAK 58

RESULT 13
 ID 0926W7 PRELIMINARY; PRT: 253 AA.
 AC 0926W7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Partition protein, Para homolog.
 GN PARA OR LMO2791 OR LIN2923.
 OS Listeria monocytogenes, and
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 NCBI_TaxID=1639, 1642;
 [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-L.monocytogenes, and L.innocua;
 RC STRAIN-BD-E / SEROVAR 1/2A, AND CLIP 11262 / SEROVAR 6A;
 RX MEDLINE-21537279; PubMed-11679669;
 RA Glasier P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
 RA Bequerio F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
 RA Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Ertlen K.-D., Fsih H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurepali G.,
 RA Meduno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
 RT "Comparative genomics of *Listeria* species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL591984; CAD01004.1; -;
 DR EMBL: AL596174; CAC98148.1; -;
 DR LISTLIST: LIN02923; -;
 DR LISTLIST: LMO02791; -;
 DR InterPro: IPR000707; ATPase_Para.
 DR Pfam: PF00991; Para; 1.
 KW Complete proteome.
 SQ SEQUENCE 253 AA; 27470 MW; 70F307B95FF33AC CRC64;

Query Match 100.0%; Score 26; DB 16; Length 253;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 Db 231 LYDAK 235

RESULT 14

09PJ15
 ID 09PJ15 PRELIMINARY; PRT: 259 AA.
 AC 09PJ15;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Periplasmic protein.
 GN C10111.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 NCBI_TaxID=197;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCTC 11168;
 RX MEDLINE-20150912; PubMed-10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham S., Holtroyd S.,
 RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrell B.G.;
 RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL: AL139074; CAB72595.1; -;
 KW Complete proteome.
 SQ SEQUENCE 259 AA; 29294 MW; AC401B4070B1D36F CRC64;

Query Match 100.0%; Score 26; DB 16; Length 259;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 Db 218 LYDAK 222

RESULT 15
 ID 09P7L6 PRELIMINARY; PRT: 281 AA.
 AC 09P7L6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative methyltransferase.
 GN SPB21C3.07C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NCBI_TaxID=4896;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H-;
 RA Seeger K., Harris D., Lyne M., Rajandream M.A., Barrell B.G.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL157918; CAB76043.1; -;
 DR InterPro: IPR001601; SAM_bind.
 DR InterPro: IPR000051; SAM_bind.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 281 AA; 32732 MW; 222B60A6AA44B039 CRC64;

Query Match 100.0%; Score 26; DB 3; Length 281;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
 Db 164 LYDAK 168

Search completed: July 11, 2003, 11:59:05
 Job time : 5.18295 secs

PT New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets
 XX
 PS Claim 1: Page 20; 47pp; English.
 XX
 CC Human MDF (see AAR48379) can be isolated from a culture of human
 CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
 CC stimulates differentiation of megakaryocytes from myeloid cells
 CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
 CC making it useful for treatment of diseases involving a decrease
 CC in platelet number (esp. thrombocytopenia) such as occurs in bone
 CC marrow transplantation and in chemotherapy. MDF has mol.wt.
 CC 55-57kd by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
 CC contains an amino acid sequence comprising at least one of the
 CC sequences AAR57107-R57115.
 CC
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 28; DB 15; Length 5;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NTEMK 5
 DB 1 NTEMK 5
 RESULT 2
 AAG89325
 ID AAG89325 standard; Protein; 163 AA.
 XX
 AC AAG89325;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 445.
 XX
 KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KW GENSET.
 XX
 OS Homo sapiens.
 XX
 PN WO200142451-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 07-DEC-2000; 2000MO-IB01938.
 XX
 PR 08-DEC-1999; 9905-0169629.
 PR 06-MAR-2000; 200005-0187470.
 XX
 PA (GENSET) GENSET.
 XX
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
 DR WPI: 2001-367870/38.
 DR N-PSDB; AAR64928.
 XX
 PT Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 21; Page 903; 921pp; English.
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patients own production of GENSET polypeptides. Conversely,

CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET polypeptide of the invention.
 CC
 SQ Sequence 163 AA;
 Query Match 100.0%; Score 28; DB 22; Length 163;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NTEMK 5
 DB 159 NTEMK 163
 RESULT 3
 AAR48379
 ID AAR48379 standard; Protein; 380 AA.
 XX
 AC AAR48379;
 XX
 DT 16-AUG-1994 (first entry)
 XX
 DE Human megakaryocyte differentiation factor.
 XX
 KW Human megakaryocyte differentiation factor; MDF; thrombopoietin;
 KW haematopoietic stimulating factor; thrombocytopenia; platelet;
 KW bone marrow transplantation; cancer chemotherapy.
 XX
 OS Homo sapiens.
 XX
 PN EP583884-A.
 XX
 PD 23-FEB-1994.
 XX
 PF 19-JUL-1993; 93EP-0305654.
 XX
 PR 17-JUL-1992; 92JP-0212305.
 PR 04-MAR-1993; 93JP-0067339.
 XX
 PA (SUNR) SUNTORI LTD.
 PA (TSUJ/) TSUJIMOTO M.
 XX
 PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
 PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamaichi K;
 XX
 DR WPI: 1994-058782/08.
 DR N-PSDB; AAO56670.
 XX
 PT New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets
 XX
 PS Claim 7; Page 30-32; 47pp; English.
 CC Human MDF can be isolated from a culture of human epidermoid
 CC carcinoma A431 cells in protein-free medium. The MDF stimulates
 CC differentiation of megakaryocytes from myeloid cells in the presence
 CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
 CC for treatment of diseases involving a decrease in platelet number
 CC (esp. thrombocytopenia) such as occurs in bone marrow
 CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kd by
 CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
 CC
 SQ Sequence 380 AA;

Query Match 100.0%; Score 28; DB 15; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
 |||||
 DB 284 NTEMK 288

RESULT 4

AA08254
 ID AAY08254 standard; Protein; 380 AA.

AC AAY08254;

DT 14-JUL-1999 (first entry)

DE Human megalin protein.

KW Megalin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 human; rat; murine.

OS Homo sapiens.

PN WO915652-A1.

PD 01-APR-1999.

PF 22-SEP-1998; 98WO-JP04269.

PR 22-SEP-1997; 97JP-0275302.

PA (KURO/) KUROKAWA K.

PA (MIYA/) MIYATA T.

PI Miyata T;

DR WPI: 1999-276983/23.

DR N-PSDB; AAX56712.

PT Megalin protein expressed specifically in mesangial cells

PS Claim 1; Page 62-64; 100pp; Japanese.

CC This invention describes the isolation of novel megalin nucleic acid and

CC proteins from human, rat and mouse tissue. This protein is expressed

CC specifically in mesangial cells. The products of the invention are

CC useful for the treatment and diagnosis of diseases involving mesangial

CC cells, such as Iga nephropathy.

SO Sequence 380 AA;

OY 1 NTEMK 5
 |||||
 DB 284 NTEMK 288

RESULT 5

AA08254
 ID AAB24142 standard; Protein; 380 AA.

AC AAB24142;

DT 30-JAN-2001 (first entry)

DE Human megalin protein sequence SEQ ID NO:2.

KW Megalin; mesangium-predominant gene; serpin regulated; nephropathy;

Iga; immunoglobulin A; detection; renal function; renal disorder;

KW diagnosis; biological sample; blood; urine.

OS Homo sapiens.

PN WO200057189-A1.

PD 28-SEP-2000.

PF 17-MAR-2000; 2000WO-JP01646.

PR 19-MAR-1999; 99JP-0075305.

PR 28-OCT-1999; 99JP-0306623.

PA (KURO/) KUROKAWA K.

PA (FUSO) FUSO PHARM IND LTD.

PA (MIYA/) MIYATA T.

PI Miyata T;

DR WPI: 2000-611642/58.

DR N-PSDB; AAX99294.

PT Evaluating renal function comprises assaying megalin protein in

PT biological sample

PS Example 2; Page 66-69; 93pp; Japanese.

CC The present invention describes a method for evaluating renal function.

CC The method comprises assaying megalin protein in biological sample. Also

CC described are: (1) use of a anti-megalin protein antibody for diagnosing

CC renal function; and (2) a kit for detecting megalin protein comprising:

CC (a) anti-megalin protein antibody attached to solid magnetic particles;

CC (b) direct or indirect fixing for the antibody to the particles; and

CC (c) a magnet. The process is useful for evaluating renal function and

CC diagnosing renal disorders by assaying megalin protein in biological

CC samples (preferably urine or blood). The process is reproducible and

CC gives accurate results. The present sequence represents the human megalin

CC protein, which is given in the exemplification of the present invention.

SO Sequence 380 AA;

OY 1 NTEMK 5
 |||||
 DB 284 NTEMK 288

Query Match 100.0%; Score 28; DB 21; Length 380;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC AAB83075;

DT 10-JUL-2001 (first entry)

DE Human megalin protein.

KW Human; megalin; mesangial cell proliferative nephritis; nephrotropic;

transgenic mouse; glomerular disease; animal model; drug screening.

OS Homo sapiens.

PN WO200124628-A1.

PD 12-APR-2001.

PF 06-OCT-2000; 2000WO-JP06988.

PR 06-OCT-1999; 99JP-0285736.

PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
XX WPI: 2001-300136/31.
DR N-PSDB: AAF82438.
XX
XX Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments -
XX
PS Example 4; Page 44-46; 62pp; Japanese.
XX
CC The present sequence is human megin. The human megin coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 28; DB 22; Length 380;
Best Local Similarity 100.0%; Pred.No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEMK 5
DB 284 NEMK 288

RESULT 7
AAG31168
ID AAG31168 standard; Protein; 428 AA.
XX
AC AAG31168;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37385.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.

```

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

```

```

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 28; DB 21; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 393 NYEMK 397

RESULT 8
AAy08215
ID AAy08215 standard; Protein; 434 AA.
XX
AC AAy08215;
XX
DT 09-JUL-1999 (first entry)
XX
DE Staphylococcus haemolyticus FemA protein.
XX
KW FemA; identification; detection; therapy; infection; femB;
KM amplification; genotyping; gram-positive bacteria; vaccine.
XX
OS Staphylococcus haemolyticus.
XX
FH Key Location/Qualifiers
FT Misc-difference 421 /note= "in frame stop codon"
FT Misc-difference 425 /note= "in frame stop codon"
FT FT
PN WO916780-A2.
XX
PD 08-APR-1999.
XX
PF 28-SEP-1998; 98WO-BE00141.
XX
PR 26-SEP-1997; 97EP-0870146.
XX
PA (BENA-) BELGIAN MIN NAT DEFENCE.
PI (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
XX Gala J, Vannuffel P;
XX WPI: 1999-287521/24.
XX DR N-PSDB; AAX37798.
XX
XX New Staphylococcus-specific oligonucleotides
XX Claim 16; Fig 6a-b; 48pp; English.
XX
XX This invention describes novel Staphylococcus-specific oligonucleotides
XX based on the consensus femA nucleotide sequence which are used to
XX develop products for the identification, detection and therapy of
XX infections. The oligonucleotides can be used for the genetic
XX amplification, the identification and/or quantification of various femA

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 5; Conservative	100.0%;	28;	100.0%;	434;
	Pred. No. 1.8e+02;			
	Mismatches 0;	Indels 0;	Gaps 0	
Db	1 NYEMK 5			
	29 NYEMK 33			
RESULT 9				
AG31167				
ID	AG31167 standard; Protein; 441 AA.			
AC	AG31167;			
XX				
DT	17-OCT-2000 (first entry)			
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 37384.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway; hydridisation assay; genetic mapping; gene expression control; promoter; termination sequence.			
XX				
OS	Arabidopsis thaliana.			
XX				
PN	EP1033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
PR	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	23-MAR-1999; 99US-0125788.			
PR	25-MAR-1999; 99US-0126264.			
PR	29-MAR-1999; 99US-0126785.			
PR	01-APR-1999; 99US-0127462.			
PR	06-APR-1999; 99US-0128234.			
PR	08-APR-1999; 99US-0128714.			
PR	16-APR-1999; 99US-0129845.			
PR	19-APR-1999; 99US-0130077.			
PR	21-APR-1999; 99US-0130449.			
PR	23-APR-1999; 99US-0130510.			
PR	23-APR-1999; 99US-0130891.			
PR	28-APR-1999; 99US-0131449.			
PR	30-APR-1999; 99US-0132048.			
PR	04-MAY-1999; 99US-0132407.			
PR	04-MAY-1999; 99US-0132484.			
PR	05-MAY-1999; 99US-0132485.			
PR	06-MAY-1999; 99US-0132486.			
PR	06-MAY-1999; 99US-0132487.			
PR	07-MAY-1999; 99US-0132863.			
PR	11-MAY-1999; 99US-0134256.			
PR	14-MAY-1999; 99US-0134218.			
PR	14-MAY-1999; 99US-0134219.			
PR	14-MAY-1999; 99US-0134221.			
PR	18-MAY-1999; 99US-0134370.			
PR	18-MAY-1999; 99US-0134768.			
PR	19-MAY-1999; 99US-0134941.			
PR	20-MAY-1999; 99US-0135124.			

PR	21-MAY-1999	9905-0135355
PR	22-MAY-1999	9905-0135659
PR	23-MAY-1999	9905-0136021
PR	27-MAY-1999	9905-0136782
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137558
PR	04-JUN-1999	9905-0137702
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	14-JUN-1999	9905-0138847
PR	16-JUN-1999	9905-0139152
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139753
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139899
PR	22-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	28-JUN-1999	9905-0140891
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141824
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142820
PR	12-JUL-1999	9905-0142877
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143524
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144384
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144614
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145211
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145518
PR	27-JUL-1999	9905-0145519
PR	28-JUL-1999	9905-0145561
PR	02-AUG-1999	9905-0146386

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 28; DB 21; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
Db 406 NYEMK 410

RESULT 10
AAG31166
ID AAG31166 standard; Protein; 479 AA.
XX
AC AAG31166;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37383.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 28; DB 21; Length 479;
Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
11111

Db 444 NYEMK 448

RESULT 11

AAV22237
ID AAV22237 standard; Protein: 847 AA.

AC AAV22237;

DT 20-SEP-1999 (first entry)

DE Human KDR signal transduction inducer protein sequence.

XX KDR signal transduction inducer protein; human; diabetic retinopathy;
KM vascular endothelial cell growth receptor; abnormal neovascularisation;
KM kinase insert domain-containing receptor; solid tumour proliferation;
KM gene therapy; metastasis; chronic rheumatoid arthritis; psoriasis;
KM retinopathy; retinopathy of prematurity.

OS Homo sapiens.

PN MO9931238-A1.

PD 24-JUN-1999.

PF 11-DEC-1998; 98MO-JP05612.

PR 12-DEC-1997; 97JP-0343474.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PA (SHIB/) SHIBUYA M.

PI Shibuya M, Yabana N;

DR WPI: 1999-405033/34.

DR N-PSDB: AAX84667.

PT KDR signal transduction inducing protein and antibodies to it

PS Claim 1; Page 68-74; 82pp; Japanese.

XX This sequence is the protein of the invention, which induces signal
CC transduction of the vascular endothelial cell growth receptor KDR (kinase
CC insert domain-containing receptor) by binding to its intracellular
CC domain. The protein can be used in the investigation, diagnosis and
CC treatment (including gene therapy) of diseases in which abnormal
CC neovascularisation takes place, such as solid tumour proliferation,
CC metastasis, chronic rheumatoid arthritis, psoriasis and retinopathy
CC (including diabetic retinopathy and retinopathy of prematurity). It may
CC be used as a screen for candidate KDR signal transduction inhibitors for
CC therapeutic use.

SO Sequence 847 AA;

Query Match 100.0%; Score 28; DB 20; Length 847;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
11111

Db 438 NYEMK 442

RESULT 12

AAAB10667
ID AAB10667 standard; Protein: 2057 AA.

AC AAB10667;

XX

DT 19-JAN-2001 (first entry)

XX L. mesenteroides alternan sucrose protein.

DE Alternan sucrose; glucosyltransferase; fructose; cosmetic; foodstuff;

KW syrup.

XX Leuconostoc mesenteroides.

OS DEL9905069-A1.

PN 10-AUG-2000.

PD 08-FEB-1999; 99DE-1005069.

PE 08-FEB-1999; 99DE-1005069.

PR (PLAN-) PLANTTEC BIOTECHNOLOGIE GMBH.

PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX Kossmann J, Welsh T, Quanz M, Knuth K;

PI WPI: 2000-550294/51.

DR N-PSDB: AAA97904.

PT New nucleic acid encoding recombinant Leuconostoc mesenteroides

PS Claim 1a; Page 30-36; 64pp; German.

XX This invention describes a novel nucleic acid molecule (I) encoding an
CC alternan sucrose (E.C. 2.4.1.140 - an enzyme, that belongs to the
CC glucosyltransferase group) The recombinant, purified alternan sucrose
CC gene is useful for the fermentative production of alternan (a
CC carbohydrate) and/or fructose by secreting the enzyme into a
CC saccharose-containing culture medium. Alternatively, the enzyme is
CC contacted with a saccharose-containing solution. The alternan and/or
CC fructose is then isolated from the medium. Cosmetic products or
CC foodstuffs containing alternan can be produced. Recombinant production of
CC alternan sucrose is advantageous as it provides a cost effective means of
CC producing fructose for high fructose containing syrups, production of
CC which previously has been achieved by costly production from maize
CC starch. This sequence represents the Leuconostoc mesenteroides alternan
CC sucrose protein which is described in the method of the invention.

SO Sequence 2057 AA;

Query Match 100.0%; Score 28; DB 21; Length 2057;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
11111

Db 1003 NYEMK 1007

RESULT 13

AAAB61548
ID AAB61548 standard; Peptide: 32 AA.

AC AAB61548;

DT 03-APR-2001 (first entry)

DE Peptide WINZIPAL used to identify hetero-associating peptides.

XX Hetero-associating coiled-coil peptide; heterodimerisation.

OS Unidentified.

PN MO200100814-A2.

XX

OY 1 NYEMK 5
111:1
Db 15 NYELK 19

Search completed: July 11, 2003, 11:54:25
Job time : 3.86902 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 0.893971 Seconds
(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-8

Perfect score: 28

Sequence: 1 NYEMK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	5	2 US-08-472-659-8	Sequence 8, Appl1
2	28	100.0	5	2 US-08-474-661-8	Sequence 8, Appl1
3	28	100.0	5	2 US-08-611-977-8	Sequence 8, Appl1
4	28	100.0	380	2 US-08-472-659-34	Sequence 34, Appl1
5	28	100.0	380	2 US-08-474-661-34	Sequence 34, Appl1
6	28	100.0	380	2 US-08-611-977-34	Sequence 34, Appl1
7	25	89.3	13	1 US-08-306-231-6	Sequence 6, Appl1
8	25	89.3	13	1 US-08-355-888A-22	Sequence 22, Appl1
9	25	89.3	13	3 US-08-693-697-22	Sequence 22, Appl1
10	25	89.3	99	3 US-08-939-696-22	Sequence 22, Appl1
11	25	89.3	235	1 US-08-946-329A-89	Sequence 89, Appl1
12	25	89.3	235	1 US-08-928-443-4	Sequence 4, Appl1
13	25	89.3	235	3 US-09-129-055-4	Sequence 4, Appl1
14	25	89.3	315	1 US-07-757-390-8	Sequence 8, Appl1
15	25	89.3	315	1 US-08-442-282-8	Sequence 8, Appl1
16	25	89.3	315	1 US-08-442-282-8	Sequence 8, Appl1
17	25	89.3	315	2 US-08-939-727-8	Sequence 8, Appl1
18	25	89.3	332	1 US-07-757-390-7	Sequence 7, Appl1
19	25	89.3	332	1 US-08-442-282-7	Sequence 7, Appl1
20	25	89.3	332	1 US-08-442-281-7	Sequence 7, Appl1
21	25	89.3	332	2 US-08-939-727-7	Sequence 7, Appl1
22	25	89.3	335	1 US-07-947-130-3	Sequence 3, Appl1
23	25	89.3	335	1 US-08-421-822-3	Sequence 3, Appl1
24	25	89.3	335	1 US-08-421-823-3	Sequence 3, Appl1
25	25	89.3	338	1 US-07-757-390-6	Sequence 6, Appl1
26	25	89.3	338	1 US-08-442-282-6	Sequence 6, Appl1
27	25	89.3	398	1 US-08-442-281-6	Sequence 6, Appl1

28	25	89.3	398	2 US-08-939-727-6	Sequence 6, Appl1
29	25	89.3	415	1 US-07-757-390-5	Sequence 5, Appl1
30	25	89.3	415	1 US-08-442-282-5	Sequence 5, Appl1
31	25	89.3	415	1 US-08-442-281-5	Sequence 5, Appl1
32	25	89.3	415	2 US-08-939-727-5	Sequence 5, Appl1
33	25	89.3	419	1 US-08-330-154-2	Sequence 2, Appl1
34	25	89.3	422	4 US-09-134-001C-5230	Sequence 5230, Ap
35	25	89.3	429	3 US-08-307-896-4	Sequence 4, Appl1
36	25	89.3	429	5 PCT-US95-11808-4	Sequence 4, Appl1
37	25	89.3	452	2 US-08-416-870C-2	Sequence 2, Appl1
38	25	89.3	534	4 US-09-321-276-4	Sequence 4, Appl1
39	25	89.3	554	4 US-08-916-481-2	Sequence 4, Appl1
40	25	89.3	563	4 US-08-916-481-3	Sequence 3, Appl1
41	25	89.3	1090	3 US-08-307-896-3	Sequence 3, Appl1
42	25	89.3	1090	3 US-08-726-214-4	Sequence 4, Appl1
43	25	89.3	1090	5 PCT-US95-11808-3	Sequence 3, Appl1
44	25	89.3	2254	2 US-08-577-010-3	Sequence 3, Appl1
45	25	89.3	2254	2 US-08-790-519-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-472-659-8
Sequence 8, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUCHI, No. 5831030uo
APPLICANT: NAKAZONO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhitro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030cm1
TITLE OF INVENTION: MECAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-472-659-8

Query Match 100.0%; Score 28; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
11111
DB 1 NYEMK 5

RESULT 2

US-08-474-661-8
Sequence 8, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253oml
TITLE OF INVENTION: MEKANARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 2213-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-661-8

Query Match 100.0%; Score 28; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
11111
DB 1 NYEMK 5

RESULT 3

US-08-611-977-8
Sequence 8, Application US/08611977
Patent No. 5972886

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886oml
TITLE OF INVENTION: MEKANARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 2213-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-8

Query Match 100.0%; Score 28; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
11111
DB 1 NYEMK 5

RESULT 4
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030cm1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 4-213305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 28; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. NO. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 284 NYEMK 288

RESULT 5
US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253cm1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-213305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANKE
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 28; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. NO. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 284 NYEMK 288

RESULT 6
US-08-611-977-34
Sequence 34, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886cm1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: P.O. Box 1404
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22131-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/611,977
;; FILING DATE: 06-MAR-1996
;; CLASSIFICATION: 435
;;
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;;
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;;
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-204
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2051
;;
;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 380 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
;; US-08-611-977-34

Query Match 100.0%; Score 28; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 284 NYEMK 288

RESULT 7
US-08-306-231-6
; Sequence 6, Application US/08306231
; Patent No. 5643748
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. R.
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan W.
; TITLE OF INVENTION: Hu-B1, 219, A NOVEL HEMATOPOIETIN
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/306,231
;; FILING DATE: 14-SEP-1994
;; CLASSIFICATION: 435
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 7225-076
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-9741/8864
;;
;; TELEEX: 66141 PENNIE
;;
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 13 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;;
;; US-08-306-231-6

Query Match 89.3%; Score 25; DB 1; Length 13;
Best Local Similarity 80.0%; Pred. No. 3.9;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 5 NYEMK 9

RESULT 8
US-08-355-888A-22
; Sequence 22, Application US/08355888A
; Patent No. 5763211
; GENERAL INFORMATION:
; APPLICANT: Snodgrass, H. R.
; APPLICANT: Cioffi, Joseph
; APPLICANT: Zupancic, Thomas J.
; APPLICANT: Shafer, Alan W.
; TITLE OF INVENTION: Hu-B1, 219, A NOVEL HUMAN HEMATOPOIETIN
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/355,888A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 7225-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide

US-08-355-888A-22

Query Match 89.3%; Score 25; DB 1; Length 13;
Best Local Similarity 80.0%; Pred. No. 3.9;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
111:1
Db 5 NYELK 9

RESULT 9

US-08-693-697-22
; Sequence 22, Application US/08693697
; Patent No. 5869610

GENERAL INFORMATION:

APPLICANT: Snodgrass, H. R.

APPLICANT: Cioffi, Joseph

APPLICANT: Zupancic, Thomas J.

TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN

TITLE OF INVENTION: RECEPTOR

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/693,697

FILING DATE: 05-AUG-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.

REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-0037-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-493-4935

TELEFAX: 650-493-5556

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-693-697-22

OY 1 NYEMK 5
111:1
Db 5 NYELK 9

RESULT 10

US-08-693-696-22
; Sequence 22, Application US/08693696
; Patent No. 6005080

GENERAL INFORMATION:

APPLICANT: Snodgrass, H. R.

APPLICANT: Cioffi, Joseph

APPLICANT: Zupancic, Thomas J.

APPLICANT: Shafer, Alan W.
TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN

TITLE OF INVENTION: RECEPTOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/693,696

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/355,888

FILING DATE: 14-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.

REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 7225-078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

OY 1 NYEMK 5
111:1
Db 5 NYELK 9

RESULT 11

US-08-946-329A-89
; Sequence 89, Application US/08946329A
; Patent No. 6057091

GENERAL INFORMATION:

APPLICANT: Beachy, Philip A.

TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: Windows 95

SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/946,329A

FILING DATE: 07-OCT-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-946-329A-89

Query Match 89.3%; Score 25; DB 3; Length 99;
Best Local Similarity 80.0%; Pred. No. 37;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYEMK 5
Db 26 NYELK 30

RESULT 12
US-08-928-443-4
Sequence 4, Application US/08928443
Patent No. 5795724
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Tang, Y. Tom
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN N-ACETYL TRANSFERASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,443
FILING DATE: HEREWITH
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0361 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 517485
US-08-928-443-4

Query Match 89.3%; Score 25; DB 1; Length 235;
Best Local Similarity 80.0%; Pred. No. 96;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 25 NYQMK 29

RESULT 13
US-09-129-055-4
Sequence 4, Application US/09129055
Patent No. 6017744

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Tang, Y. Tom
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN N-ACETYL TRANSFERASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/129,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/928,443
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0361 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 517485
US-09-129-055-4

Query Match 89.3%; Score 25; DB 3; Length 235;
Best Local Similarity 80.0%; Pred. No. 96;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 25 NYQMK 29

RESULT 14
US-07-757-390-8
Sequence 8, Application US/07757390
Patent No. 5453491
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tomioka, Akira
APPLICANT: Takegi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/757,390
FILING DATE: 19910910
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: AMINO ACID
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-757-390-8
Query Match 89.3%; Score 25; DB 1; Length 315;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTYEMK 5
DB 251 NTYELK 255
RESULT 15
US-08-442-282-8
Sequence 8, Application US/08442282
Patent No. 5760204
GENERAL INFORMATION:
APPLICANT: Takatsu, Kiyoshi
APPLICANT: Tomioka, Akira
APPLICANT: Takegi, Satoshi
APPLICANT: Murata, Yoshiyuki
TITLE OF INVENTION: Human And Murine Interleukin-5 Receptor
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,282
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/757,390
FILING DATE: 10-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7005-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 8698864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-442-282-8
Query Match 89.3%; Score 25; DB 1; Length 315;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTYEMK 5
DB 251 NTYELK 255
Search completed: July 11, 2003, 12:02:11
Job time: 1.89397 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 1.4657 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-8

Perfect score: 28

Sequence: 1 NEMK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	5	9 US-10-091-442-8	Sequence 8, Appl1
2	28	100.0	5	9 US-09-140-719-8	Sequence 8, Appl1
3	28	100.0	163	10 US-09-731-872-445	Sequence 445, App
4	28	100.0	380	9 US-10-091-442-34	Sequence 34, Appl
5	28	100.0	380	10 US-09-140-719-34	Sequence 34, Appl
6	28	100.0	399	9 US-10-029-180-64	Sequence 12, Appl
7	28	100.0	584	9 US-09-995-749A-12	Sequence 12, Appl
8	28	100.0	847	9 US-10-177-293-498	Sequence 498, App
9	25	89.3	299	10 US-09-861-451A-16	Sequence 36, Appl
10	25	89.3	446	10 US-09-853-386-69	Sequence 69, Appl
11	25	89.3	554	9 US-09-878-672-2	Sequence 2, Appl1
12	25	89.3	554	10 US-09-800-336-4	Sequence 4, Appl1
13	25	89.3	563	9 US-09-878-672-3	Sequence 3, Appl1
14	25	89.3	625	10 US-09-853-386-63	Sequence 63, Appl
15	25	89.3	626	10 US-09-853-386-64	Sequence 64, Appl
16	25	89.3	626	10 US-09-853-386-65	Sequence 65, Appl
17	25	89.3	626	10 US-09-853-386-96	Sequence 96, Appl
18	25	89.3	643	10 US-09-853-386-70	Sequence 70, Appl
19	24	85.7	42	10 US-09-864-761-42476	Sequence 42476, A

20	24	85.7	376	10 US-09-815-242-5669	Sequence 5669, Ap
21	24	85.7	377	10 US-09-815-242-12700	Sequence 12700, A
22	24	85.7	377	10 US-09-815-242-12726	Sequence 12726, A
23	24	85.7	557	9 US-10-081-872-206	Sequence 206, App
24	24	85.7	635	10 US-09-841-132-520	Sequence 520, App
25	24	85.7	896	9 US-10-210-296-5	Sequence 5, Appl1
26	24	85.7	898	9 US-10-270-875-37	Sequence 37, Appl
27	24	85.7	898	9 US-10-270-878-37	Sequence 37, Appl
28	24	85.7	898	9 US-10-270-786-37	Sequence 37, Appl
29	24	85.7	898	9 US-10-270-710-37	Sequence 37, Appl
30	24	85.7	898	9 US-10-270-859-37	Sequence 37, Appl
31	23	82.1	5	10 US-09-832-312-67	Sequence 67, Appl
32	23	82.1	67	10 US-09-864-761-33978	Sequence 33978, A
33	23	82.1	76	9 US-10-156-761-15015	Sequence 15015, A
34	23	82.1	83	9 US-10-091-572-237	Sequence 237, App
35	23	82.1	83	9 US-09-764-891-3348	Sequence 3348, Ap
36	23	82.1	110	10 US-09-815-242-13253	Sequence 13253, A
37	23	82.1	122	10 US-09-867-550-1862	Sequence 1862, Ap
38	23	82.1	132	9 US-10-093-766-44	Sequence 44, Appl
39	23	82.1	132	9 US-10-171-311-89	Sequence 89, Appl
40	23	82.1	132	9 US-10-197-666A-70	Sequence 70, Appl
41	23	82.1	132	9 US-10-197-666A-146	Sequence 146, App
42	23	82.1	132	9 US-10-197-666A-148	Sequence 148, App
43	23	82.1	133	9 US-09-854-133-124	Sequence 124, App
44	23	82.1	133	9 US-10-197-666A-72	Sequence 72, Appl
45	23	82.1	133	9 US-10-197-666A-74	Sequence 74, Appl

ALIGNMENTS

RESULT 1
US-10-091-442-8
Sequence 8, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSOJIMOTO, Masafumi
IMASA, Fuyuki
TSUBOBUOKA, No. US20020164711A1uo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, No. US20020164711A1uhlro
KURIHARA, Tatsuya
YAMACHIKI, Kozo
YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140, 719
FILING DATE: 08-Aug-1998
APPLICATION NUMBER: US 08/474, 661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339

FILED DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
TELEPHONE: (703) 836-2021
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-091-442-8

Query Match 100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 1 NYEMK 5

RESULT 2
US-09-140-719-8
Sequence 8, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUBOKAWA, Hiroshi
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1om1
TITLE OF INVENTION: MEKANARYCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-8

Query Match 100.0%; Score 28; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 1 NYEMK 5

RESULT 3
US-09-731-872-445
Sequence 445, Application US/09731872
Patent No. US20020102604A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78, US, REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 445
LENGTH: 163
TYPE: PRT
ORGANISM: Homo sapiens
US-09-731-872-445

Query Match 100.0%; Score 28; DB 10; Length 163;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 159 NYEMK 163

RESULT 4
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUBOKAWA, Hiroshi
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1om1
TITLE OF INVENTION: MEKANARYCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

ADDRESSER: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-Aug-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34
Query Match 100.0%; Score 28; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NYEMK 5
Db 284 NYEMK 288

RESULT 5
US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUTSUMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. US20010026931AUO
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931AUH1RO
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931AUOM1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States

ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-Aug-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34
Query Match 100.0%; Score 28; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NYEMK 5
Db 284 NYEMK 288

RESULT 6
US-10-029-180-64
Sequence 64, Application US/10029180
Publication No. US20020182708A1
GENERAL INFORMATION:
APPLICANT: Call, Brian M.
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin T.
APPLICANT: Milna, G. Todd
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeffery C.
APPLICANT: Trueheart, Josh
TITLE OF INVENTION: No. US20020182708A1el Regulators of Fungal Gene Expression
FILE REFERENCE: MIC-004
CURRENT APPLICATION NUMBER: US/10/029,180
PRIOR FILING DATE: 2001-12-22
PRIOR APPLICATION NUMBER: US 60/257,431
NUMBER OF SEQ ID NOS: 138
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 64
LENGTH: 399
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: fungal gene

US-10-029-180-64

Query Match 100.0%; Score 28; DB 9; Length 399;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
|||||
DB 384 NYEMK 388

RESULT 7

US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 100.0%; Score 28; DB 9; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
|||||
DB 413 NYEMK 417

RESULT 8

US-10-177-293-498
; Sequence 498, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatf, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Cannavarpu, Manjula
; APPLICANT: Kamalkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, VIC
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzstai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21

; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 498
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-498

Query Match 100.0%; Score 28; DB 9; Length 847;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
|||||
DB 438 NYEMK 442

RESULT 9
US-09-861-451A-36
; Sequence 36, Application US/09861451A
; Patent No. US20020068289A1
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: FF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PP7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Deduced protein
US-09-861-451A-36

Query Match 89.3%; Score 25; DB 10; Length 299;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
|||||
DB 117 NYEMK 121

RESULT 10
US-09-853-386-69
; Sequence 69, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnahan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURF
; FILE REFERENCE: P019720S1

;; CURRENT APPLICATION NUMBER: US/09/853,386
;; CURRENT FILING DATE: 2001-05-11
;; PRIOR APPLICATION NUMBER: US 60/203645
;; PRIOR FILING DATE: 2000-05-12
;; NUMBER OF SEQ ID NOS: 153
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 69
;; LENGTH: 446
;; TYPE: PR
;; ORGANISM: Sus scrofa
US-09-853-386-69

Query Match 89.3%; Score 25; DB 10; Length 446;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
Db 68 NYELK 72

RESULT 11
US-09-878-672-2
;; Sequence 2, Application US/09878672
;; Publication No. US20030049812A1
;; GENERAL INFORMATION:
;; APPLICANT: Wallis, Nicola G.
;; TITLE OF INVENTION: NOVEL histidine kinase
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dechert Price & Rhoads
;; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: US
;; ZIP: 19103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/878,672
;; FILING DATE: 11-Jun-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/916,481
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dickinson, Todd O
;; REGISTRATION NUMBER: 28,354
;; REFERENCE/DOCKET NUMBER: GM10022-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-994-2252
;; TELEFAX: 215-994-2222
;; TELEX: <Unknown>
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 554 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-878-672-2

Query Match 89.3%; Score 25; DB 9; Length 554;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
Db 529 NYOMK 533

RESULT 12
US-09-800-396-4
;; Sequence 4, Application US/09800396
;; Patent No. US20020065395A1
;; GENERAL INFORMATION:
;; APPLICANT: Wallis, Nicola
;; TITLE OF INVENTION: NOVEL RESPONSE REGULATOR
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dechert Price & Rhoads
;; STREET: 997 Lenox Drive, Building 3, Suite 210
;; CITY: Lawrenceville
;; STATE: NJ
;; COUNTRY: USA
;; ZIP: 08543

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/800,396
;; FILING DATE: 06-Mar-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/879,531
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bloom, Allen
;; REGISTRATION NUMBER: 29,135
;; REFERENCE/DOCKET NUMBER: GM10018
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 609-520-3214
;; TELEFAX: 609-520-3259
;; TELEX: <Unknown>
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 554 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-800-396-4

Query Match 89.3%; Score 25; DB 10; Length 554;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
Db 529 NYOMK 533

RESULT 13
US-09-878-672-3
;; Sequence 3, Application US/09878672
;; Publication No. US20030049812A1
;; GENERAL INFORMATION:
;; APPLICANT: Wallis, Nicola G.
;; TITLE OF INVENTION: NOVEL histidine kinase
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dechert Price & Rhoads
;; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: US
;; ZIP: 19103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: DOS

```

; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/878,672
;   FILING DATE: 11-Jun-2001
;   CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/916,481
;   FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Dickinson, Todd O
;   REGISTRATION NUMBER: 28,354
;   REFERENCE/DOCKET NUMBER: GM10022-1
;
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 215-994-2252
;   TELEFAX: 215-994-2222
;   TELEX: <Unknown>
;
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 563 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
;
; US-09-878-672-3
;
Query Match      89.3%; Score 25; DB 9; Length 563;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 NYEMK 5
Db      538 NYOMK 542

RESULT 14
US-09-853-386-63
; Sequence 63, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
;   APPLICANT: Murphy, Evelyn
;   APPLICANT: Bresnihan, Barry
;   APPLICANT: Conneely, Orla
;   APPLICANT: Fitzgerald, Oliver
;   TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
;   FILE REFERENCE: P019720U1
;   CURRENT APPLICATION NUMBER: US/09/853,386
;   CURRENT FILING DATE: 2001-05-11
;   PRIOR APPLICATION NUMBER: US 60/203645
;   PRIOR FILING DATE: 2000-05-12
;   NUMBER OF SEQ ID NOS: 153
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 63
;     LENGTH: 625
;     TYPE: PRT
;     ORGANISM: HUMAN
;
US-09-853-386-63

Query Match      89.3%; Score 25; DB 10; Length 625;
Best Local Similarity 80.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 NYEMK 5
Db      68 NYELK 72

RESULT 15
US-09-853-386-64
; Sequence 64, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
;   APPLICANT: Murphy, Evelyn
```

```

; APPLICANT: Bresnihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
;   TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
;   FILE REFERENCE: P019720U1
;   CURRENT APPLICATION NUMBER: US/09/853,386
;   CURRENT FILING DATE: 2001-05-11
;   PRIOR APPLICATION NUMBER: US 60/203645
;   PRIOR FILING DATE: 2000-05-12
;   NUMBER OF SEQ ID NOS: 153
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 64
;     LENGTH: 626
;     TYPE: PRT
;     ORGANISM: HUMAN
;
US-09-853-386-64

Query Match      89.3%; Score 25; DB 10; Length 626;
Best Local Similarity 80.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 NYEMK 5
Db      68 NYELK 72

Search completed: July 11, 2003, 12:37:34
Job time : 1.4657 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 1.00832 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-8

Perfect score: 28
Sequence: 1 NYEMK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	124	1 MN21B	nonstructural prot
2	28	100.0	124	1 MN21B	nonstructural prot
3	28	100.0	128	2 A49166	3',5'-cyclic-AMP p
4	28	100.0	330	2 D97235	probable hydrolase
5	28	100.0	358	2 T25790	hypothetical prote
6	28	100.0	420	2 B97276	glycosyltransferas
7	28	100.0	479	2 T09365	cytochrome P450 ho
8	28	100.0	538	2 S49933	hypothetical prote
9	28	100.0	608	2 T18445	hypothetical prote
10	28	100.0	900	2 T14277	myosin-like protei
11	28	100.0	2059	2 T41933	large tegument pro
12	25	89.3	62	2 T36921	gene MHC DQ-alpha
13	25	89.3	62	2 T61802	gene MHC DQ-alpha
14	25	89.3	75	2 S40739	hypothetical prote
15	25	89.3	133	2 S12097	rfai protein - Sal
16	25	89.3	182	2 T33023	hypothetical prote
17	25	89.3	193	1 S39401	hydrogenase matur
18	25	89.3	214	2 S45465	heat shock protein
19	25	89.3	220	2 S64784	hypothetical prote
20	25	89.3	235	2 T38333	ARD1 N-acetyl tran
21	25	89.3	265	2 G49508	integrase/recombin
22	25	89.3	267	2 S43900	methyl coenzyme M
23	25	89.3	289	2 E97185	polysaccharide dea
24	25	89.3	300	2 S55620	capsid protein 26
25	25	89.3	300	2 T33232	hypothetical prote
26	25	89.3	335	2 H91191	hypothetical prote
27	25	89.3	335	2 A86039	probable LPS biosy
28	25	89.3	337	2 AH0972	lipopolysaccharide
29	25	89.3	339	2 E83419	probable permease

30	25	89.3	377	2 B64428	formate hydrogenly
31	25	89.3	386	2 H90485	glucose 1-dehydrog
32	25	89.3	406	2 S64493	hypothetical prote
33	25	89.3	411	2 C64501	probable phosphono
34	25	89.3	415	2 S12357	interleukin-5 rece
35	25	89.3	416	2 T16058	hypothetical prote
36	25	89.3	416	2 S19896	plasmidogen activa
37	25	89.3	422	2 JC5325	methicillin resist
38	25	89.3	438	2 B83295	hypothetical prote
39	25	89.3	476	2 B82955	adenylate cyclase
40	25	89.3	502	2 I37136	lipoprotein IplA -
41	25	89.3	523	2 F86217	protein T2767.12 I
42	25	89.3	563	2 C97944	histidine kinase (
43	25	89.3	567	2 T34339	hypothetical prote
44	25	89.3	625	2 S71930	neuron-derived rec
45	25	89.3	664	1 S73624	hypothetical prote

ALIGNMENTS

RESULT 1
MN21B
nonstructural protein 1B - human respiratory syncytial virus
C:Species: human respiratory syncytial virus
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: B94336; C93010; A04032
R:Collins, P.L.; Wertz, G.W.
Virology 143, 442-451, 1985
A:Title: Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human r
A:Reference number: A94336; MUID:86045905; PMID:2998021
A:Accession: B94336
A:Molecule type: genomic RNA
A:Residues: 1-124 <COL>
A:Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:K03349; GB:M11217;
1.1: PID:9333927
R:Elango, N.; Satake, M.; Venkatesan, S.
J. Virol. 55, 101-110, 1985
A:Title: mRNA sequence of three respiratory syncytial virus genes encoding two nonstr
A:Reference number: A93010; MUID:85237684; PMID:4009789
A:Accession: C93010
A:Molecule type: genomic RNA
A:Residues: 1-54, 'R', 56-124 <ELA>
A:Cross-references: GB:M11486
C:Genetics:
A:Gene: 1B
C:Superfamily: respiratory syncytial virus nonstructural protein 1B
C:Keywords: nonstructural protein

Query Match 100.0%; Score 28; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYEMK 5
Db 64 NYEMK 68

RESULT 2
MN21B
nonstructural protein 1B - human respiratory syncytial virus (strain 16537)
N:Alternate names: nonstructural protein NS2
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: B32063
R:Johnson, P.R.; Collins, P.L.
J. Gen. Virol. 70, 1539-1547, 1989
A:Title: The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus (C
A:Reference number: A32063; MUID:89279331; PMID:2525176
A:Accession: B32063
A:Molecule type: mRNA
A:Residues: 1-124 <JOH>
A:Cross-references: EMBL:DD0736; NID:9222559; PIDD:BAA00636.1; PID:9222561

C:Genetics:
A:Gene: 1B
C:Superfamily: respiratory syncytial virus nonstructural protein 1B
C:Keywords: nonstructural protein

Query Match 100.0%; Score 28; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 64 NTEMK 68

RESULT 3

A49166
3',5'-cyclic-AMP phosphodiesterase (EC 3.1.4.-), cyclic-GMP inhibited - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 19-May-2000
C:Accession: A49166
R:Lebon, T.R.; Kasuya, J.; Paxton, R.J.; Belfrage, P.; Hockman, S.; Manganiello, V.C.; F
Endocrinology 130, 3265-3274, 1992
A:Title: Purification and characterization of guanosine 3',5'- monophosphate-inhibited 1
A:Reference number: A49166; MUID:92283180; PMID:1317779
A:Accession: A49166
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-128 <EB>
C:Superfamily: cyclic-nucleotide phosphodiesterase, CGMP-inhibited; 3',5'-cyclic-nucleo
C:Keywords: CGMP binding; phosphoric diester hydrolase

Query Match 100.0%; Score 28; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 17 NTEMK 21

RESULT 4

D97235
Probable hydrolase from alpha/beta family, YOKD B. subtilis ortholog [Imported] - Clostr
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: D97235
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK80671.1; PID:g15025760; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2725

Query Match 100.0%; Score 28; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 209 NTEMK 213

RESULT 5

T25790
hypothetical protein F53E10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25790
R:Beck, C.; Mamsley, P.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid F53E10.
A:Reference number: Z20087
A:Accession: T25790
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-358 <BEC>
A:Cross-references: EMBL:U88177; PIDN:AA842291.1; GSPDB:GN00023; CESP:F53E10.5
A:Experimental source: strain Bristol N2; clone F53E10
C:Genetics:
A:Gene: CESP:F53E10.5
A:Map position: 5
A:introns: 23/3; 51/3; 73/1; 103/1; 122/1; 178/1; 191/3; 221/3; 317/3

Query Match 100.0%; Score 28; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 47 NTEMK 51

RESULT 6

B97276
glycosyltransferase [Imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97276
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97276
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-420 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK80997.1; PID:g15026118; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3057

Query Match 100.0%; Score 28; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 191 NTEMK 195

RESULT 7

T09365
cytochrome P450 homolog F23K16.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 16-Feb-2001
C:Accession: T09365
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16652
A:Accession: T09365
A:Molecule type: DNA
A:Residues: 1-479 <BEV>
A:Cross-references: EMBL:AL078620; GSPDB:GN00062; ATSP:F23K16.120
A:Experimental source: cultivar Columbia; BAC clone F23K16
C:Genetics:
A:Gene: ATSP:F23K16.120
A:Map position: 4
C:Superfamily: Candida cytochrome P450 52A3; cytochrome P450 homology

F:267-446/Domain: cytochrome P450 homology <P45>

Query Match 100.0%; Score 28; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 444 NYEMK 448

RESULT 8

S49933 hypothetical protein YIL045W - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YI9905.03

C:Species: Saccharomyces cerevisiae

C>Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 06-Feb-1998

C:Accession: S49933

R:Odeall, C.; Bowman, S.

submitted to the EMBL Data Library, December 1994

A:Reference number: S49931

A:Accession: S49933

A:Molecule type: DNA

A:Residues: 1-538 <ODE>

A:Cross-references: GB:Z47047; EMBL:Z46861; NID:g603997; PID:g763301; MIPS:YIL045W

C:Genetics:

A:Gene: SGD:PIG2

A:Cross-references: SGD:S0001307; MIPS:YIL045W

A:Map position: 9L

C:Superfamily: glucamylase starch-binding domain homology

F:396-516/Domain: glucamylase starch-binding domain homology <SBD>

Query Match 100.0%; Score 28; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 47;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 525 NYEMK 529

RESULT 9

T18445

hypothetical protein C0420W - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2002

C:Accession: T18445

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A:Accession: T18445

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-608 <LAM>

A:Cross-references: EMBL:Z98547; NID:e1325376; PIDN:CAB11118.2

C:Genetics:

A:Map position: 3

A:Introns: 326/3; 490/3; 533/3

A:Note: C0420W

Query Match 100.0%; Score 28; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 53;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 270 NYEMK 274

RESULT 10

T14277

myosin-like protein my3 - common sunflower

N:Alternate names: unconventional myosin

C:Species: Helianthus annuus (common sunflower)

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 23-Mar-2001

C:Accession: T14277

R:Yugrek, O.

submitted to the EMBL Data Library, March 1997

A:Description: Molecular cloning of a new mini myosin from sunflower with similaritie

A:Reference number: Z14226

A:Accession: T14277

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-900 <VUG>

A:Cross-references: EMBL:U94783; NID:g2731701; PID:g2731702

C:Genetics:

A:Gene: MY3

C:Superfamily: myosin motor domain homology

F:168-807/Domain: myosin motor domain homology <MMO>

Query Match 100.0%; Score 28; DB 2; Length 900;
Best Local Similarity 100.0%; Pred. No. 80;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 280 NYEMK 284

RESULT 11

T41933

large tegument protein - human herpesvirus 7 (strain J1)

C:Species: human herpesvirus 7

A:Variety: strain J1

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T41933

R:Nicholas, J.

submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of huma

A:Reference number: Z22022

A:Accession: T41933

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2059 <NIC>

A:Cross-references: EMBL:U43400; PIDN:AAC54693.1

A:Experimental source: strain J1

C:Genetics:

A:Note: U31

C:Superfamily: varicella-zoster virus 240K tegument protein

Query Match 100.0%; Score 28; DB 2; Length 2059;
Best Local Similarity 100.0%; Pred. No. 1,9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 1182 NYEMK 1186

RESULT 12

I36921

gene MHC DQ-alpha 1 protein - hamadryas baboon (fragment)

C:Species: Papio hamadryas (hamadryas baboon)

C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999

C:Accession: I36921

R:Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.

Immunogenetics 36, 71-78, 1992

A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.

A:Reference number: I36899; MUID:92307745; PMID:1612647

A:Accession: I36921

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-62 <RES>

A:Cross-references: GB:M76226; NID:g176597; PIDN:AAA5391.1; PID:g176598

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 89.3%; Score 25; DB 2; Length 62;
 Best Local Similarity 80.0%; Pred. No. 28;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYEMK 5
 ||:|
 Db 56 NYQMK 60

RESULT 13

161802
 gene MHC DQ-alpha 1 protein - rhesus macaque (fragment)
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 23-Jul-1999
 C:Accession: I61802
 R:Kenter, M.; Oetting, N.; Anholts, J.; Leunissen, J.A.M.; Jonker, M.; Bontrop, R.E.
 Immunogenetics 36, 71-78, 1992
 A:Title: Evolutionary relationships among the primate MHC-DQA1 and DQA2 alleles.
 A:Reference number: I36899; MUID:92307745; PMID:1612647
 A:Accession: I61802
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-62 <RRS>
 A:Cross-references: GB:M76229; NID:9342161; PIDN:AAA36870.1; PID:9342162
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

89.3%; Score 25; DB 2; Length 62;
 Best Local Similarity 80.0%; Pred. No. 28;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
 ||:|

Db 56 NYQMK 60

RESULT 14

S40739
 hypothetical protein ZK507.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
 C:Accession: S40739
 R:Hawkins, T.; Thomas, K.
 submitted to the EMBL Data Library, December 1993
 A:Reference number: S40735
 A:Accession: S40739
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-75 <HAM>
 A:Cross-references: EMBL:Z29116; NID:g1067087; PID:g439269
 C:Genetics:
 A:introns: 32/1

Query Match

89.3%; Score 25; DB 2; Length 75;
 Best Local Similarity 80.0%; Pred. No. 35;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
 ||:|

Db 47 NYEMK 51

RESULT 15

S12097
 rfaI protein - Salmonella typhimurium
 C:Species: Salmonella typhimurium
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
 C:Accession: S12097
 R:Carstenius, P.; Flock, J.I.; Lindberg, A.
 Nucleic Acids Res. 18, 6128, 1990
 A:Title: Nucleotide sequence of rfaI and rfaJ genes encoding lipopolysaccharide glycosyl
 A:Reference number: S12097; MUID:91045080; PMID:2235496
 A:Accession: S12097
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-133 <CAR>
 A:Cross-references: EMBL:X53847; NID:g47882; PIDN:CAA37841.1; PID:g47883

Query Match

89.3%; Score 25; DB 2; Length 133;
 Best Local Similarity 80.0%; Pred. No. 62;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
 ||:|

Db 41 NYELK 45

Search completed: July 11, 2003, 12:00:42
 Job time : 2.10832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 11, 2003, 11:43:41 ; Search time 0.47817 Seconds

(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-8

Perfect score: 28

Sequence: 1 NYEMK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt_40.*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	124	VNS2_BRSVA	065695 bovine resp
2	28	100.0	124	VNS2_HRSVA	P24569 human resp
3	28	100.0	124	VNS2_HRSVA	P04543 human resp
4	28	100.0	124	VNS2_HRSVB	042038 human resp
5	28	100.0	124	VNS2_HRSVL	086305 human resp
6	28	100.0	124	VNS2_ORSVW	065707 ovine resp
7	28	100.0	380	SPB7_HUMAN	075635 homo sapien
8	28	100.0	538	PIG2_YEAST	P40187 saccharomyc
9	28	100.0	847	VAV3_HUMAN	091068 homo sapien
10	28	100.0	847	VAV3_MOUSE	P52362 human herpe
11	28	100.0	2059	TEGU_HSV7J	093525 lactobacill
12	25	89.3	149	UREE_LACFE	045251 bradyrhizob
13	25	89.3	193	HUPD_BRAJA	P36416 dictyostel
14	25	89.3	203	AROH_DICDI	P15992 saccharomyc
15	25	89.3	213	HS26_YEAST	P41227 homo sapien
16	25	89.3	235	AROH_HUMAN	P19816 salmonella
17	25	89.3	337	FEAT_SALTY	P53288 saccharomyc
18	25	89.3	406	YG43_YEAST	Q59007 methanococc
19	25	89.3	411	APG1_METJA	P21183 mus musculu
20	25	89.3	415	ILSR_MOUSE	P29524 rattus norv
21	25	89.3	416	PAIR_RAT	P37966 bacillus su
22	25	89.3	502	LPLA_BACSU	092570 homo sapien
23	25	89.3	626	NRA3_HUMAN	F75234 mycoplasma
24	25	89.3	664	Y366_MYCPN	067012 aquifex aeo
25	25	89.3	666	SPOT_AOUAE	P22258 acetogenium
26	25	89.3	762	SLAP_ACEKI	006708 saccharomyc
27	25	89.3	880	YL86_YEAST	008462 homo sapien
28	25	89.3	887	CYAA_HUMAN	092927 rickettsia
29	25	89.3	906	SECA_RICPR	P26769 rattus norv
30	25	89.3	1090	CYAA_RAT	P16419 gallus gall
31	25	89.3	1131	MYPF_CHICK	Q10178 schistosach
32	25	89.3	1188	S3B1_SCHPO	P06274 marchantia
33	25	89.3	1386	RPDD_MARPO	

ALIGNMENTS

RESULT 1	VNS2_BRSVA	STANDARD	PRT	124 AA.	
AC	065695				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	Nonstructural protein 2 (Nonstructural protein 1B).				
GN	1B OR NS2.				
OS	Bovine respiratory syncytial virus (strain A51908) (BRS).				
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;				
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.				
OX	NCBI_TaxID=11247;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=95146950; PubMed=7844532;				
RA	Pasteur M.K., Samal S.K.;				
RT	"Nucleotide sequence analysis of the non-structural NS1 (1C) and NS2 (1B) protein genes of bovine respiratory syncytial virus.";				
RL	J. Gen. Virol. 76:193-197(1995).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL: U15938; AAA85672.1; -				
DR	InterPro: IPR004336; RSV_NS2.				
DR	Pfam: PF03113; RSV_NS2.1.				
KW	Nonstructural protein.				
SO	SEQUENCE 124 AA; 14576 MW; B47548B0D5FB8FFD CRC64;				
QY	1 NYEMK 5				
DB	64 NYEMK 68				
Query Match	100.0%; Score 28; DB 1; Length 124;				
Best Local Similarity	100.0%; Pred. No. 6.1;				
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
RESULT 2	VNS2_HRSVA	STANDARD	PRT	124 AA.	
ID	VNS2_HRSVA				
AC	P24569				
DT	01-MAR-1992 (Rel. 21, Created)				
DT	01-MAR-1992 (Rel. 21, Last sequence update)				
DT	01-MAR-1992 (Rel. 21, Last annotation update)				
DE	Nonstructural protein 2 (Nonstructural protein 1B).				
GN	1B OR NS2.				
OS	Human respiratory syncytial virus (subgroup B / strain 18537).				
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;				
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.				

0929K9 bacillus ha
P33012 escherichia
015145 homo sapien
088593 mus musculu
057640 methanococc
P12758 escherichia
033808 salmonella
008444 klebsiella
057629 methanococc
P43770 haemophilus
049402 mycoplasma
Q9v1h6 pyrococcus

```

OX NCBI_TaxID-11251;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89279331; PubMed-2525176;
RA Johnson P.R., Collins P.L.;
RT "The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial
RT virus (RSV) of antigenic subgroups A and B: sequence conservation and
RT divergence within RSV genomic RNA.";
RL J. Gen. Virol. 70:1539-1547(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00736; BAA00636.1; -
DR PIR: B32063; MNZ215;
DR InterPro: IPR004336; RSV_NS2.
DR Pfam: PF03113; RSV_NS2; 1.
DR Nonstructural protein.
KW SEQUENCE 124 AA; 14624 MW; 526F8C2C6553C1DA CRC64;

Query Match 100.0%; Score 28; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 64 NYEMK 68

RESULT 3
VNS2_HRSVA
ID VNS2_HRSVA STANDARD; PRT; 124 AA.
AC P04543;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural protein 2 (Nonstructural protein 1B).
GN 1B OR NS2.
OS Human respiratory syncytial virus (strain A2).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID-11259;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86045905; PubMed-2998021;
RA Collins P.L., Wertz G.W.;
RT "Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of
RT human respiratory syncytial virus.";
RL Virology 143:442-451(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-85237684; PubMed-4009789;
RA Elango N., Satake M., Venkatesan S.;
RT "mRNA sequence of three respiratory syncytial virus genes encoding
RT two nonstructural proteins and a 22k structural protein.";
RL J. Virol. 55:101-110(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-95266253; PubMed-7747420;
RA Connors M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;
RT "A cold-passaged, attenuated strain of human respiratory syncytial
RT virus contains mutations in the F and L genes.";
RL Virology 208:478-484(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE-97187925; PubMed-9035372;
RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,
RA Murphy B.R.;

```

```

RT "Acquisition of the ts phenotype by a chemically mutagenized cold-
RT passed human respiratory syncytial virus vaccine candidate results
RT from the acquisition of a single mutation in the polymerase (L)
RT gene.";
RL Virus Genes 13:269-273(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M11486; AAB59851.1; -
DR EMBL: U50362; AAB86657.1; -
DR EMBL: U50363; AAB86667.1; -
DR EMBL: U63644; AAC55963.1; -
DR PIR: A04032; MNZ21B;
DR InterPro: IPR004336; RSV_NS2.
DR Pfam: PF03113; RSV_NS2; 1.
DR Nonstructural protein.
FT CONFLICT 55 R -> K (IN REF. 1).
KW SEQUENCE 124 AA; 14702 MW; C14DF0BE9C9E5512 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 64 NYEMK 68

RESULT 4
VNS2_HRSVB
ID VNS2_HRSVB STANDARD; PRT; 124 AA.
AC Q42038;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural protein 2 (Nonstructural protein 1B).
GN 1B OR NS2.
OS Human respiratory syncytial virus (strain B1).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID-79692;
RN [1]
RP SEQUENCE FROM N.A.
RA Karon R.A., Buonagurio D.A., Georgiu A.F., Whitehead S.S.,
RA Adamus J.E., Clements-Mann M.L., Harris D.O., Randolph V.B.,
RA Udem S.A., Murphy B.R., Sidhu M.S.;
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF013254; AAB82430.1; -
DR EMBL: AF013255; AAB82441.1; -
DR InterPro: IPR004336; RSV_NS2.
DR Pfam: PF03113; RSV_NS2; 1.
DR Nonstructural protein.
KW SEQUENCE 124 AA; 14567 MW; A9FD37D964F97073 CRC64;

```

```

Oy      1 NFEK 5
        |||||
Db      64 NFEK 68

RESULT 5
VNS2_HRSVL STANDARD; PRT; 124 AA.
ID VNS2_HRSVL
AC 086305;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Nonstructural protein 2 (Nonstructural protein 1B).
GN 1B OR NS2.
OS Human respiratory syncytial virus (subgroup A / strain Long).
OC Viruses: ssRNA negative-strand viruses; Mononegavirales.
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11260;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U35029; AAA79090.1; -
DR InterPro: IPR004336; RSV_NS2.
DR Pfam: PF03113; RSV_NS2; 1.
KM Nonstructural protein.
SQ SEQUENCE 124 AA; 14705 MW; 98657338FEB744E1 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 NFEK 5
        |||||
Db      64 NFEK 68

RESULT 6
VNS2_ORSV STANDARD; PRT; 124 AA.
ID VNS2_ORSV
AC 065707;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUL-1998 (Rel. 37, Last annotation update)
DE Nonstructural protein 2 (Nonstructural protein 1B).
GN 1B.
OS Ovine respiratory syncytial virus (strain MSU 83-1578) (ORSV).
OC Viruses: ssRNA negative-strand viruses; Mononegavirales.
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=19699;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94157498; PubMed-8113762;
RA Alenquer H.M., Potgieter L.N.D.;
RT "Nucleotide and predicted amino acid sequence analysis of the ovine
RT respiratory syncytial virus non-structural 1C and 1B genes and the
RT small hydrophobic protein gene."
RL J. Gen. Virol. 75:401-404(1994).
RN -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L15451; AAA42813.1; -
DR InterPro: IPR004336; RSV_NS2.
DR Pfam: PF03113; RSV_NS2; 1.
KM Nonstructural protein.
SQ SEQUENCE 124 AA; 14496 MW; 201D0F5DFEA9A4EB CRC64;

Query Match 100.0%; Score 28; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 NFEK 5
        |||||
Db      64 NFEK 68

RESULT 7
SPB7_HUMAN STANDARD; PRT; 380 AA.
ID SPB7_HUMAN
AC 075635;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Megsin (TF55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-97326116; PubMed-9182567;
RA Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.;
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity."
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE-Mesangial cells.
RX MEDLINE-98376492; PubMed-9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, megin, is a new serpin upregulated in
RT IgA nephropathy."
RL J. Clin. Invest. 102:828-836(1998).
RN -----
CC -!- FUNCTION: Might function as an inhibitor of Lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D88575; BAA31232.1; -
DR EMBL: AF027866; AAC64506.1; -
DR HSSP: P05619; 1HLE.
DR Genev: HGNC:13902; SERPINB7.
DR MIM: 603557; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.

```

DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 347 REACTIVE_BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFF605 CRC64;
 Query Match 100.0%; Score 28; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NTEMK 5
 DB 284 NTEMK 288

RESULT 8
 PIG2_YEAST STANDARD; PRT; 538 AA.
 AC P40187;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE GSY2 interacting protein PIG2.
 GN PIG2 OR YII045W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Bartell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Genes S., Hamlyn N., Horsnell T.S., Hunt S., Jagals K., Jones M.,
 RA Louis E., Lye G., Moule S., Mouton R., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RA Submitted (Dec-1994) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN PARTIAL SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RP MEDLINE-97197965; PubMed-9046081;
 RA Cheng C., Huang D., Roach P.J.;
 RT "Yeast PIG genes: PIG1 encodes a putative type 1 phosphatase subunit
 RT that interacts with the yeast glycogen synthase Gsy2p.";
 RL Yeast 13:1-8(1997).
 CC -1- FUNCTION: INTERACTS WITH GLYCOGEN SYNTHASE 2 (GSY2); POSSIBLY ALSO
 CC INTERACTS WITH PHOSPHATASE 1 (GLC7).
 CC -1- SIMILARITY: TO YEAST GIP2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage -by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: 246861; CAA86906.1; -;
 DR SGD: S0001307; PIG2.
 DR InterPro: IPR005036; CBM_21.
 DR Pfam: PF03370; CBM_21; 1.
 SQ SEQUENCE 538 AA; 61938 MW; 73ED8F5CC6692172 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 538;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 5; Conservative 0; Mismatches -0; Indels 0; Gaps 0;
 OY 1 NTEMK 5
 DB 525 NTEMK 529

RESULT 9
 VAV3_HUMAN

ID VAV3_HUMAN STANDARD; PRT; 847 AA.
 AC Q9UKW4; Q95230; Q9Y5X8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vav-3 protein.
 GN VAV3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE-99455043; PubMed-10523675;
 RA Movilla N., Bustelo X.R.;
 RT "Biological and regulatory properties of Vav-3, a new member of the
 RT Vav family of oncoproteins."
 RL Mol. Cell. Biol. 19:7870-7885(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Breast, and Colon carcinoma;
 RX MEDLINE-98371222; PubMed-9705494;
 RA Trenkle T., Welsh J., Jung B., Mathieu-Daude F., McClelland M.;
 RT "Non-stoichiometric reduced complexity probes for cDNA arrays."
 RL Nucleic Acids Res. 26:3883-3891(1998).
 CC -1- FUNCTION: EXCHANGE FACTOR FOR GTP-BINDING PROTEINS RHOA, RHOG AND,
 CC TO A LESSER EXTENT, RAC-1. BINDS PHYSICALLY TO THE NUCLEOTIDE-FREE
 CC STATES OF THOSE GTPASES.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, ALPHA (SHOWN HERE) AND BETA;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 DEL-HOMOLOGY (DH) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage -by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF118887; AAD20349.1; -;
 DR EMBL: AF118886; AAD20348.1; -;
 DR EMBL: AF067817; AAC79695.1; -;
 DR HSSP: P29355; ISEM.
 DR Genew; HGNC:12659; VAV3.
 DR MIM: 605541; -;
 DR InterPro: IPR003247; CH_type.
 DR InterPro: IPR001715; Calponin-like.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR001311; GDS_CDC24.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000219; RhogEF.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR003096; SM22_calponin.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00130; DAG_PE-bind; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00307; CH; 1.
 DR Pfam: PF00621; RhogEF; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRINTS: PR00888; SM22CALPONIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 1.
 DR ProDom: PD001527; CH_type; 1.

DR SMART; SM00109; C1; 1.
 DR SMART; SM00033; CH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RHOGEF; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS50021; CH; 1.
 DR PROSITE; PS500479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50010; DH_2; 1.
 DR PROSITE; PS00741; DH_1; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 2.
 DR Phorbol-ester binding; zinc; SH2 domain; SH3 domain; Repeat;
 KM Guanine-nucleotide releasing factor; Alternative splicing.
 FT DOMAIN 1 119
 FT DOMAIN 192 371
 FT DOMAIN 400 502
 FT DOMAIN 514 562
 FT DOMAIN 592 660
 FT DOMAIN 672 766
 FT DOMAIN 788 847
 FT VARSPLIC 1 107
 FT MEPMKQCAQMLHCKVLPTRNRYTMDSAQVFDLAQTLRDGV
 FT LCOLLNLRRAHSINKETLRPOMQSCFLKNIREFLTC
 FT CEEFGMRKSELEAFDLFDVDRDGK -> MQLPDCQRAHL
 FT P (IN ISOFORM BETA).
 FT K -> E (IN REF. 2).
 FT Y -> H (IN REF. 1); AAD20348).
 FT T -> S (IN REF. 2).
 FT V -> A (IN REF. 1); AAD20348).
 FT CONFLICT 217 217
 FT CONFLICT 217 217
 FT CONFLICT 298 298
 FT CONFLICT 429 429
 SQ SEQUENCE 847 AA; 97775 MW; C1E29F0B094CB721 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 847;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 438 NYEMK 442
 RESULT 10
 VAV3_MOUSE
 ID VAV3_MOUSE STANDARD; PRT; 847 AA.
 AC 09ROC8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vav-3 protein.
 GN VAV3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 ON NCBI_Taxid=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20179693; PubMed=10713454;
 RA Tremlle T., McClelland M., Adlkofer K., Welsh J.;
 RT "Major transcript variants of VAV3, a new member of the VAV family of
 RT guanine nucleotide exchange factors.";
 RL Gene 245:119-149(2000).
 CC -1- FUNCTION: EXCHANGE FACTOR FOR GTP-BINDING PROTEINS RHOA, RHOG AND,
 CC TO A LESSER EXTENT, RAC-1. BINDS PHYSICALLY TO THE NUCLEOTIDE-FREE
 CC STATES OF THOSE GTPASES (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 CALPONIN-HOMOLOGY (CH) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 CC BINDING DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF067816; AAF09171.1; -.
 DR HSSP; P29355; ISEM.
 DR MGD; MGI:1888518; Vav3.
 DR InterPro; IPR003247; CH type.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR001331; GDS_CDC24.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000219; RHOGEF.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR003096; SM22-calponin.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00130; DAG_PE-bind; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00307; CH; 1.
 DR Pfam; PF00621; RHOGEF; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRINTS; PR00888; SM22CALPONIN.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 1.
 DR ProDom; PD001527; CH type; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00033; CH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00325; RHOGEF; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS50021; CH; 1.
 DR PROSITE; PS500479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50010; DH_2; 1.
 DR PROSITE; PS500741; DH_1; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 2.
 KM Phorbol-ester binding; zinc; SH2 domain; SH3 domain; Repeat;
 KW Guanine-nucleotide releasing factor.
 FT DOMAIN 1 119
 FT DOMAIN 192 371
 FT DOMAIN 400 502
 FT DOMAIN 514 562
 FT DOMAIN 592 660
 FT DOMAIN 672 766
 FT DOMAIN 788 847
 SQ SEQUENCE 847 AA; 97946 MW; 9A6B63F0D9E60F8F CRC64;
 Query Match 100.0%; Score 28; DB 1; Length 847;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 438 NYEMK 442
 RESULT 11
 TEGU_HSV7J
 ID TEGU_HSV7J STANDARD; PRT; 2059 AA.
 AC P52362;
 DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large tegument protein.
 GN U31.
 OS Human herpesvirus (type 7 / strain J1) (HHV7).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 NCBI_TaxID=57278;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Nicholas J.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TEGUMENT PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
 CC EHV-1 24, EBV BPLF1, HSV-1 64, VZV 22, AND HCMV UL48.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U03400; AAC54693.1;
 SQ SEQUENCE 2059 AA; 239476 MW; AAE6CA6DEDC5D316 CRC64;
 Query Match 100.0%; Score 28; DB 1; Length 2059;
 Best Local Similarity 100.0%; Pred. No. 1; Leu02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NTEMK 5
 DB 1182 NTEMK 1186
 RESULT 12
 UREB_LACFE
 ID UREB_LACFE STANDARD; PRT; 149 AA.
 AC 09X525;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Urease accessory protein uree.
 GN UREB.
 OS Lactobacillus fermentum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 NCBI_TaxID=1613;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN-TK1214;
 RA Colon E., Visser J.J., Van Vuuren H.J.J.;
 RL "Urease operon of Lactobacillus fermentum."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN UREASE METALLOCENTER ASSEMBLY. BINDS NICKEL.
 CC PROBABLY FUNCTIONS AS A NICKEL DONOR DURING METALLOCENTER ASSEMBLY
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE UREB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF120718; AAD22481.1;
 DR InterPro: IPR004029; Uree.
 DR Pfam: PF02814; Uree; 1.
 KM Nickel.

SQ SEQUENCE 149 AA; 17041 MW; 419F74BC0BF4A14 CRC64;
 Query Match 89.3%; Score 25; DB 1; Length 149;
 Best Local Similarity 80.0%; Pred. No. 39;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NTEMK 5
 DB 130 NTEMK 134
 RESULT 13
 HUPD_BRAJA
 ID HUPD_BRAJA STANDARD; PRT; 193 AA.
 AC 045251; 045249;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hydrogenase expression/formation protein hupd.
 GN HUPD.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bradyrhizobium group; Bradyrhizobium.
 NCBI_TaxID=375;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94047099; PubMed-8230232;
 RA Van Soom C., Browneys J., Verreth C., Vanderleyden J.;
 RT "Nucleotide sequence analysis of four genes, hupc, hupd, hupf and
 RT hupg, downstream of the hydrogenase structural genes in
 RT Bradyrhizobium japonicum."
 RL J. Mol. Biol. 234:508-512(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-USDA 110;
 RX MEDLINE-94215887; PubMed-8163174;
 RA Fu C., Maier R.J.;
 RT "Sequence and characterization of three genes within the hydrogenase
 RT gene cluster of Bradyrhizobium japonicum."
 RL Gene 141:47-52(1994).
 CC -1- FUNCTION: NOT KNOWN. COULD BE INVOLVED IN THE PROCESSING OF
 CC HYDROGENASE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M52.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z21948; CAA79944.1;
 DR EMBL: L24446; AAD13469.1;
 DR HSSP: P37182; ICF2.
 DR MEROPS: M52.001;
 DR InterPro: IPR004419; HupD.
 DR InterPro: IPR000671; Hydrgn_uptake.
 DR Pfam: PF01750; Hycl; 1
 DR PRINTS: PR00446; HYDRGNUPAKE.
 DR TIGRPAWS: TIGR00072; hydrgn_prot; 1.
 DR TIGRPAWS: TIGR00140; hupd; 1.
 KW HydroLase; Metalloprotease; Nickel.
 FT METAL 23
 FT METAL 69
 FT METAL 100
 FT METAL 149
 FT CONFLICT 149 176 A -> S (IN REF. 2).
 FT CONFLICT 171 176 IWRSPS -> LAESE (IN REF. 2).
 SQ SEQUENCE 193 AA; 21429 MW; 34A8A2B53176ADAD CRC64;
 Query Match 89.3%; Score 25; DB 1; Length 193;
 Best Local Similarity 80.0%; Pred. No. 50;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYEMK 5
1111:
Db 187 NYEMK 191

RESULT 14

ARND_DICDI STANDARD; PRT: 203 AA.
ID ARND_DICDI
AC P36416:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-terminal acetyltransferase complex ARD1 subunit homolog.
GN NMTA OR ARD1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RA Mueller-Taubenberger A.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SEEMS TO BE INVOLVED IN N-ACETYLTATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ACETYLTTRANSFERASE FAMILY. ARD1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U06453; AAA16510.1; -.
CC DR DictyDb; DD05073; nta.
CC DR InterPro; IPR000183; GCSAcetyltransf.
CC DR Pfam; PF00583; Acetyltransf; 1.
CC KM Transferrase; Acyltransferase.
CC SO SEQUENCE 203 AA; 23415 MW; 501B6B46DF0F146 CRC64;
Query Match 89.3%; Score 25; DB 1; Length 203;
Best Local Similarity 80.0%; Pred. No. 53;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 NYEMK 5
11:11
Db 26 NYOMK 30

RESULT 15

HS26_YEAST STANDARD; PRT: 213 AA.
ID HS26_YEAST
AC P15992;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Heat shock protein 26 (26 kDa heat shock protein).
GN HSP26 OR YBR072W OR YBR0714.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90097950; Pubmed=2689876;
RA Susek R.E., Lindquist S.L.;
RT "hsp26 of Saccharomyces cerevisiae is related to the superfamily of
RT small heat shock proteins but is without a demonstrable function."
RL Mol. Cell. Biol. 9:5265-5271(1989).
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=89378758; Pubmed=2673926;
RA Bossier P., Fitch I.T., Boucherie H., Tuile M.F.;
RT "Structure and expression of a yeast gene encoding the small
RT heat-shock protein Hsp26.";
RL Gene 78:323-330(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95076715; Pubmed=7985423;
RA van der Aart O.J.M., Barthe C., Dolignon F., Aigle M., Crouzet M.,
RA Steensma H.Y.;
RT "Sequence analysis of a 31 kb DNA fragment from the right arm of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 10:959-964(1994).
CC -1- FUNCTION: NOT KNOWN. ONE OF THE MAJOR POLYPEPTIDES PRODUCED ON
CC HEAT SHOCK.
CC -1- SUBUNIT: PRESENT IN LARGE COMPLEXES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE ENTRY INTO STATIONARY
CC PHASE RESULTING FROM GLUCOSE LIMITATION.
CC -1- INDUCTION: BY HEAT SHOCK, AND UNDER OTHER CONDITIONS OF STRESS,
CC SUCH AS INCREASED SALT CONCENTRATION AND STARVATION.
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M23871; AAA66914.1; -.
DR EMBL; M26942; AAA79010.1; -.
DR EMBL; X76294; CAA53929.1; -.
DR EMBL; Z35941; CAA85016.1; -.
DR PIR; S45465; S45465.
DR PIR; S39222; S39222.
DR SGD; S0000276; HSP26.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Heat shock.
FT INT_MET 0
FT CONFLICT 31
FT CONFLICT 206
FT CONFLICT 206
SO SEQUENCE 213 AA; 23748 MW; 1C9C4B0D0626B6A4 CRC64;

Query Match 89.3%; Score 25; DB 1; Length 213;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 NYEMK 5
11:11
Db 102 NYELK 106

Search completed: July 11, 2003, 11:55:20
Job time : 2.47817 secs

THIS PAGE BLANK (USPTO)

QY	1	NYEMK	5
Db	6	NYEMK	10
RESULT 2			
QY1859			
ID	Q91859		
TC	Q91859;	PRELIMINARY;	PRT; 124 AA

DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Nonstructural protein 2.
 GN NS2.
 OS Bovine respiratory syncytial virus.
 OC Parainflaviridae; Pneumovirinae; Pneumovirus.
 NCBI_TaxID=11246;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=88LU195;
 RA Larsen L.E., Gottschalk E., Blixenkronne-Møller M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054664; AAC3667.1; -
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 SQ SEQUENCE 124 AA; 14677 MW; 20311D79D7915EB5 CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 |||||
 DB 64 NYEMK 68

RESULT 3

O9YWM9 PRELIMINARY; PRT; 124 AA.
 AC O9YWM9;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Nonstructural protein 2.
 GN NS2.
 OS Bovine respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Parainflaviridae; Pneumovirinae; Pneumovirus.
 NCBI_TaxID=11246;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=88CVA70;
 RA Larsen L.E., Gottschalk E., Blixenkronne-Møller M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054665; AAC36670.1; -
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 SQ SEQUENCE 124 AA; 14645 MW; 20311D7B2ABE9BA0 CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 |||||
 DB 64 NYEMK 68

RESULT 4

O9YWM8 PRELIMINARY; PRT; 124 AA.
 AC O9YWM8;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Nonstructural protein 2.
 GN NS2.
 OS Bovine respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Parainflaviridae; Pneumovirinae; Pneumovirus.
 NCBI_TaxID=11246;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9304899;
 RA Larsen L.E., Gottschalk E., Blixenkronne-Møller M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054666; AAC36673.1; -
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 SQ SEQUENCE 124 AA; 14635 MW; 13311D7B29BE9AE9 CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 |||||
 DB 64 NYEMK 68

RESULT 5

O9YWM6 PRELIMINARY; PRT; 124 AA.
 AC O9YWM6;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Nonstructural protein 2.
 GN NS2.
 OS Bovine respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Parainflaviridae; Pneumovirinae; Pneumovirus.
 NCBI_TaxID=11246;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9402020;
 RA Larsen L.E., Gottschalk E., Blixenkronne-Møller M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054667; AAC36676.1; -
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 SQ SEQUENCE 124 AA; 14649 MW; 35311D219ABE8EB5 CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 |||||
 DB 64 NYEMK 68

RESULT 6

O9YWM4 PRELIMINARY; PRT; 124 AA.
 AC O9YWM4;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Nonstructural protein 2.
 GN NS2.
 OS Bovine respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Parainflaviridae; Pneumovirinae; Pneumovirus.
 NCBI_TaxID=11246;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9314893;
 RA Larsen L.E., Gottschalk E., Blixenkronne-Møller M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054668; AAC36679.1; -
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 SQ SEQUENCE 124 AA; 14605 MW; 23F9095B2ABC5C65 CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 64 NYEMK 68

RESULT 7

O9YS29 PRELIMINARY; PRT; 124 AA.
 AC O9YS29;
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Nonstructural protein 2.
 GN NS2.
 OS Bovine respiratory syncytial virus, and
 OS Bovine respiratory syncytial virus (strain A51908) (BRS).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11246, 11247;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Bovine respiratory syncytial virus; STRAIN-ATUES1908;
 RX MEDLINE-99102581; PubMed-9847328;
 RA Buchholz U.J., Flink S., Conzelmann K.K.;
 RT Generation of bovine respiratory syncytial virus (BRSV) from cDNA:
 RT BRSV NS2 is not essential for virus replication in tissue culture, and
 RT the human RSV leader region acts as a functional BRSV genome
 RT promoter.";
 RL J. Virol. 73:251-259(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Bovine respiratory syncytial virus; STRAIN-ATUES1908;
 RA Buchholz U.J., Flink S., Conzelmann K.K.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Bovine respiratory syncytial virus (strain A51908) (BRS);
 RX STRAIN-ATUES1908;
 RC MEDLINE-21580793; PubMed-11724268;
 RA Yunus A.S., Khattar S.K., Collins P.L., Samal S.K.;
 RT Rescue of bovine respiratory syncytial virus from cloned cDNA: entire
 RT genome sequence of BRSV strain A51908.";
 RL Virus Genes 23:157-164(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Bovine respiratory syncytial virus (strain A51908) (BRS);
 RC STRAIN-ATUES1908;
 RA Yunus A.S., Khattar S.K., Collins P.L., Samal S.K.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF295543; AAC96302.1; -;
 DR EMBL: AF295543; AAC96302.1; -;
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 DR SEQUENCE 124 AA; 14605 MW; D80A8317FABE99AC CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 64 NYEMK 68

RESULT 8
 ID 012405 PRELIMINARY; PRT; 124 AA.
 AC 012405;

DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Non-structural protein 2 (NS2).
 GN NS2.
 OS Human respiratory syncytial virus, and
 OS respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250, 12814;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S2 AND S2 TS1C;
 RX MEDLINE-97185152; PubMed-9032893;
 RA Tolley K.P., Marriott A.C., Simpson A., Plovs D.J., Matthews D.A.,
 RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
 RA Pringle C.R.;
 RT Identification of mutations contributing to the reduced virulence of
 RT a modified strain of respiratory syncytial virus.";
 RL Vaccine 14:1637-1646(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S2 AND S2 TS1C;
 RA Easton A.J.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U39662; AAC57021.1; -;
 DR EMBL: U39662; AAC57021.1; -;
 DR InterPro: IPR004336; RSV_NS2.
 DR Pfam: PF03113; RSV_NS2; 1.
 DR SEQUENCE 124 AA; 14729 MW; 1922830810B9B4F1 CRC64;

Query Match 100.0%; Score 28; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 64 NYEMK 68

RESULT 9

O9AN06 PRELIMINARY; PRT; 133 AA.
 AC O9AN06;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ID28.
 GN ID28.
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bradyrhizobium group; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-110SPC4;
 RX MEDLINE-21101824; PubMed-11157954;
 RA Gottfert M., Rothlisberger S., Kundig C., Beck C., Marty R.,
 RA Hennecke H.;
 RT Potential symbiosis-specific genes uncovered by sequencing a 410-kb
 RT DNA region of the Bradyrhizobium japonicum chromosome.";
 RL J. Bacteriol. 183:1405-1412(2001).
 DR EMBL: AF322012; AAG60707.1; -;
 DR HSRF: P37182; ICF2.
 DR MEROPS: M52.001; -;
 DR InterPro: IPR000671; Hydrgn_uptake.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF01750; HycI; 1.
 DR TIGRGRAMS: TIGR00072; hydrgn_prot; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR SEQUENCE 133 AA; 14605 MW; FD371219966FD2F7 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 120 NYEMK 124

RESULT 10

ID 089012 PRELIMINARY; PRT; 320 AA.

AC 089012; 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)
DE Magea8 protein.
GN Magea8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;
RX MEDLINE=9134295; PubMed=9933564;
RA De Plaen E., de Backer O., Arnaud D., Bonjean B., Chomez P.,
Martelange V., Avner P., Baldacel P., Babinet C., Hwang S.Y.,
Knowles B., Boon T.;
RT "A new family of mouse genes homologous to the human MAGE genes.";
RL Genomics 55:176-184(1999).
DR EMBL; AJ005532; CAA06586.1; -;
DR MGD; MGI:1333827; Magea8.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 320 AA; 36214 MW; 414DEDD9925565C CRC64;

Query Match 100.0%; Score 28; DB 11; Length 320;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 110 NYEMK 114

RESULT 11

ID 089010 PRELIMINARY; PRT; 325 AA.

AC 089010; 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)
DE Magea6 protein.
GN Magea6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;
RX MEDLINE=9134295; PubMed=9933564;
RA De Plaen E., de Backer O., Arnaud D., Bonjean B., Chomez P.,
Martelange V., Avner P., Baldacel P., Babinet C., Hwang S.Y.,
Knowles B., Boon T.;
RT "A new family of mouse genes homologous to the human MAGE genes.";
RL Genomics 55:176-184(1999).
DR EMBL; AJ005530; CAA06584.1; -;
DR MGD; MGI:1333837; Magea6.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR002190; MAGE.

DR Pfam; PF01454; MAGE; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ SEQUENCE 325 AA; 36640 MW; 4F582109EC03383A CRC64;

Query Match 100.0%; Score 28; DB 11; Length 325;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 110 NYEMK 114

RESULT 12

ID 097FL1 PRELIMINARY; PRT; 330 AA.

AC 097FL1; 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, last sequence update)
DE Predicted hydrolase from alpha/beta family, YQXD B.subtilis
ortholog.
GN CAC2725.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacilli; Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,
Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
Bennett G.N., Koonin E.V., Smith D.R.;
RT Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL EMBL; AE007770; AAK80671.1; -;
DR EMBL; AE007770; AAK80671.1; -;
DR InterPro; IPR000379; Ser_estr_site.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 330 AA; 37432 MW; E0C2160806B744D6 CRC64;

Query Match 100.0%; Score 28; DB 16; Length 330;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
DB 209 NYEMK 213

RESULT 13

ID P91317 PRELIMINARY; PRT; 358 AA.

AC P91317; 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)
DE F53E10.5 protein.
GN F53E10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Snelton N., Smith A., Sonhammer E., Staden R., Sultson J.,
 RA Thierri-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Beck C., Mamsley P.;
 RT "The sequence of C. elegans cosmid F53E10.";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U88177; AAB42291.1; -;
 SQ SEQUENCE 358 AA; 41249 MW; 288D3CA725629352 CRC64;

Query Match 100.0%; Score 28; DB 5; Length 358;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 47 NYEMK 51

RESULT 14

O92FC7 PRELIMINARY; PRT; 420 AA.
 AC O92FC7;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Factor essential for methicillin resistance FEMa.
 GN FEMa.
 OS Staphylococcus haemolyticus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1283;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC2970;
 RA Vannuffel P., Heusterprente M., Bouyer M., Philippe M., Gala J.-L.;
 RT "Molecular characterization of fema from Staphylococcus hominis,
 RT Staphylococcus saprophyticus and Staphylococcus haemolyticus and fema-
 RT based discrimination of staphylococcal species.";
 RL Res. Microbiol. 0:0-0(1998).
 DR EMBL: AF099962; AAC69631.1; -;
 DR InterPro: IPR003447; Meth_resist.
 DR Pfam: PF02388; FemAB; 1.
 SQ SEQUENCE 420 AA; 49582 MW; 4492B0C6598F60C2 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 420;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 29 NYEMK 33

RESULT 15

O9S6V4 PRELIMINARY; PRT; 420 AA.
 AC O9S6V4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE FEMa.
 GN FEMa.
 OS Staphylococcus haemolyticus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1283;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hoskins J., Alborn J.T.W.E., Unal S., Flokowitsch J.E., Greaney M.,
 RA Skatrud P.L.;
 RT "Cloning and Characterization of fema and femb from Staphylococcus
 RT epidermidis and Staphylococcus haemolyticus.";
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U23711; AAD2132.1; -;
 DR InterPro: IPR003447; Meth_resist.
 DR Pfam: PF02388; FemAB; 1.
 SQ SEQUENCE 420 AA; 49559 MW; 1C003811D2C59C1E CRC64;

Query Match 100.0%; Score 28; DB 2; Length 420;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
 DB 29 NYEMK 33

Search completed: July 11, 2003, 11:59:07
 Job time : 4.18295 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 5.73805 Seconds
(without alignments)
232.223 Million cell updates/sec

Title: US-10-091-442-9
Perfect score: 50
Sequence: 1 AVAMKHQERK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq-101002:*

1:	/SID2/gcgdata/geneseq/emb1/AA1980.DAT:*
2:	/SID2/gcgdata/geneseq/emb1/AA1981.DAT:*
3:	/SID2/gcgdata/geneseq/emb1/AA1982.DAT:*
4:	/SID2/gcgdata/geneseq/emb1/AA1983.DAT:*
5:	/SID2/gcgdata/geneseq/emb1/AA1984.DAT:*
6:	/SID2/gcgdata/geneseq/emb1/AA1985.DAT:*
7:	/SID2/gcgdata/geneseq/emb1/AA1986.DAT:*
8:	/SID2/gcgdata/geneseq/emb1/AA1987.DAT:*
9:	/SID2/gcgdata/geneseq/emb1/AA1988.DAT:*
10:	/SID2/gcgdata/geneseq/emb1/AA1989.DAT:*
11:	/SID2/gcgdata/geneseq/emb1/AA1990.DAT:*
12:	/SID2/gcgdata/geneseq/emb1/AA1991.DAT:*
13:	/SID2/gcgdata/geneseq/emb1/AA1992.DAT:*
14:	/SID2/gcgdata/geneseq/emb1/AA1993.DAT:*
15:	/SID2/gcgdata/geneseq/emb1/AA1994.DAT:*
16:	/SID2/gcgdata/geneseq/emb1/AA1995.DAT:*
17:	/SID2/gcgdata/geneseq/emb1/AA1996.DAT:*
18:	/SID2/gcgdata/geneseq/emb1/AA1997.DAT:*
19:	/SID2/gcgdata/geneseq/emb1/AA1998.DAT:*
20:	/SID2/gcgdata/geneseq/emb1/AA1999.DAT:*
21:	/SID2/gcgdata/geneseq/emb1/AA2000.DAT:*
22:	/SID2/gcgdata/geneseq/emb1/AA2001.DAT:*
23:	/SID2/gcgdata/geneseq/emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	380	15	AA48379
2	50	100.0	380	20	AA708254
3	50	100.0	380	21	AA24142
4	50	100.0	380	22	AA83075
5	44	88.0	10	15	AA57115
6	41	82.0	380	20	AA08255
7	41	82.0	380	21	AA24150
8	41	82.0	380	22	AA64286
9	41	82.0	380	22	AA83076
10	40	80.0	131	23	ABP34239

11	37	74.0	368	20	AA08256
12	37	74.0	368	21	AA24151
13	37	74.0	368	22	AA83077
14	36	72.0	118	22	ABG11324
15	34	68.0	67	20	AA11676
16	34	68.0	82	20	AAV3595
17	34	68.0	82	21	AA12159
18	34	68.0	83	20	AAW78170
19	34	68.0	221	20	ABG11323
20	33	66.0	226	20	AA04707
21	33	66.0	223	22	AA59373
22	33	66.0	309	22	AA678271
23	33	66.0	337	20	AAV3493
24	33	66.0	700	22	AAW23827
25	33	66.0	700	22	AAW38831
26	33	66.0	717	22	AAW40617
27	33	66.0	793	13	AAW20743
28	33	66.0	793	17	AAW02282
29	33	66.0	793	20	AAV04695
30	33	66.0	1232	22	ABW63685
31	32	64.0	65	22	AA678029
32	32	64.0	136	23	ABW90000
33	32	64.0	235	21	AAO20091
34	32	64.0	235	23	AAU97023
35	32	64.0	236	20	AAV04700
36	32	64.0	254	22	AAW59372
37	32	64.0	299	20	AAW89198
38	32	64.0	310	22	AA678270
39	32	64.0	474	22	ABW58359
40	32	64.0	581	22	ABW63013
41	32	64.0	643	22	ABG22551
42	32	64.0	708	22	AAW33073
43	32	64.0	798	15	AAW57283
44	32	64.0	802	13	AAW20744
45	32	64.0	802	17	AAW02283

ALIGNMENTS

RESULT 1
AA48379 standard; Protein; 380 AA.

AC	AA48379;
DT	16-AUG-1994 (first entry)
XX	
DE	Human megakaryocyte differentiation factor.
XX	
KW	Human megakaryocyte differentiation factor; MDF; thrombopoietin;
KW	hematopoietic stimulating factor; thrombocytopoiesis; platelet;
KW	bone marrow transplantation; cancer chemotherapy.
XX	
OS	Homo sapiens.
XX	
PN	EP583884-A.
XX	
PD	23-FEB-1994.
XX	
PF	19-JUL-1993; 93BP-0305654.
XX	
PR	17-JUL-1992; 92JP-0212305.
PR	04-MAR-1993; 93JP-0067339.
PA	(SUNR) SUNTORY LTD.
PA	(TSUJ) TSUJIMOTO M.
XX	
PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
PI	Tsujimoto M, Tsunoka N, Yamaguchi N, Yamachi K;
XX	
DR	WPI; 1994-058782/08.
DR	N-PSDB; AA056670.

Mouse megalin prote
Mouse megalin prote
Murine megalin prote
Novel human diagno
Human 5' EST seque
Extended human sec
Hydrophobic domain
Human secreted pro
Novel human diagno
Mouse RPTA amino
Human protein tyro
Human PTP-epsilon
Chlamydia pneumoni
Human EST encoded
Human polypeptide
Human polypeptide
Murine receptor-ty
Murine receptor-ty
Mouse receptor-ty
Drosophila melanog
Piscirickettsia sa
Human polypeptide
Protein encoded by
Cattle enteropepti
Human RPTA amino
Human protein tyro
Aspartate-specific
Human PTP-alpha-DI
Drosophila melanog
Drosophila melanog
Novel human diagno
Human protein segm
Bovine enterokinas
Human receptor-ty
Human receptor-ty

```

XX  New megakaryocyte differentiation factor - isolated from human
PT  epidermoid carcinoma cells, used to treat conditions involving a
XX  decrease in platelets
PS  Claim 7; Page 30-32; 47pp; English.
XX
CC  Human MDF can be isolated from a culture of human epidermoid
CC  carcinoma A431 cells in protein-free medium. The MDF stimulates
CC  differentiation of megakaryocytes from myeloid cells in the presence
CC  of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC  for treatment of diseases involving a decrease in platelet number
CC  (esp. thrombocytopenia) such as occurs in bone marrow
CC  transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by
CC  SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
XX
SQ  Sequence 380 AA;

Query Match 100.0%; Score 50; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMMHQERK 10
DB 204 AVAMMHQERK 213

RESULT 2
AA08254
ID AAY08254 standard; Protein; 380 AA.
XX
AC AAY08254;
XX
DT 14-JUL-1999 (first entry)
XX
DE Human megsin protein.
XX
KW Megsin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
KW human; rat; murine.
XX
OS Homo sapiens.
XX
PN WO915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 1999-276983/23.
DR N-PSDB: AAX56712.
XX
SQ Megsin protein expressed specifically in mesangial cells

Claim 1; Page 62-64; 100pp; Japanese.
XX
XX This invention describes the isolation of novel megsin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as IgA nephropathy.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 50; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 AVAMMHQERK 10
DB 204 AVAMMHQERK 213

RESULT 3
AAB24142
ID AAB24142 standard; Protein; 380 AA.
XX
AC AAB24142;
XX
DT 30-JAN-2001 (first entry)
XX
DE Human megsin protein sequence SEQ ID NO:2.
XX
KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
KW IgA; immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
OS Homo sapiens.
XX
PN WO200057189-A1.
XX
PD 28-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-JP01646.
XX
PR 19-MAR-1999; 99JP-0075305.
PR 28-OCT-1999; 99JP-0306823.
XX
PA (KURO/) KUROKAWA K.
PA (FUSO) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 2000-611642/58.
DR N-PSDB: AAA99294.
XX
PT Evaluating renal function comprises assaying megsin protein in
PT biological sample -
XX
PS Example 2; Page 66-69; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megsin protein in biological sample. Also
CC described are: (1) use of a anti-megsin protein antibody for diagnosing
CC renal function; and (2) a kit for detecting megsin protein comprising:
CC (a) anti-megsin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying megsin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence represents the human megsin
CC protein, which is given in the exemplification of the present invention.
XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 50; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMMHQERK 10
DB 204 AVAMMHQERK 213

RESULT 4
AAB83075
ID AAB83075 standard; Protein; 380 AA.
XX
AC AAB83075;

```


XX 10-JUL-2001 (first entry)
 DT XX
 XX
 DE Human megasin protein.
 XX
 KW Human; megasin; mesangial cell proliferative nephritis; nephrotropic;
 KM transgenic mouse; glomerular disease; animal model; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO200124628-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 06-OCT-2000; 2000WO-JP06988.
 XX
 PR 06-OCT-1999; 99JP-0285736.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2001-300136/31.
 DR N-PSDB; AAF82438.
 XX
 PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments -
 PS Example 4; Page 44-46; 62pp; Japanese.
 XX
 CC The present sequence is human megasin. The human megasin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 CC
 SQ Sequence 380 AA;
 XX
 QY Query Match 100.0%; Score 50; DB 22; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.06; Mismatches 0; Gaps 0;
 Matches 10; Conservative 0; Indels 0; Gaps 0;
 1 AVAMMHQERK 10
 DB 204 AVAMMHQERK 213

RESULT 5
 AAR57115
 ID AAR57115 standard; peptide: 10 AA.
 XX
 AC AAR57115;
 XX
 DT 16-AUG-1994 (first entry)
 XX
 DE Human megakaryocyte differentiation factor peptide 9.
 XX
 KW Human megakaryocyte differentiation factor; MPF; thrombopoietin;
 KM haematopoietic stimulating factor; thrombocytopoiesis; platelet;
 XX bone marrow transplantation; cancer chemotherapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 8 /note- "not determined"
 XX
 PN EP583884-A.

XX 23-FEB-1994.
 PD XX
 XX 19-JUL-1993; 93EP-0305654.
 FE XX
 PR 17-JUL-1992; 92JP-0212305.
 PR 04-MAR-1993; 93JP-0067339.
 XX
 PA (SUNR) SUNTORY LTD.
 PA (TSUJ/) TSUJIMOTO M.
 XX
 PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
 PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
 XX
 DR WPI: 1994-058782/08.
 XX
 PT New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets
 XX
 PS Claim 1; Page 21; 47pp; English.
 XX
 CC Human MDF (see AAR48379) can be isolated from a culture of human
 CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
 CC stimulates differentiation of megakaryocytes from myeloid cells
 CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
 CC making it useful for treatment of diseases involving a decrease
 CC in platelet number (esp. thrombocytopaenia) such as occurs in bone
 CC marrow transplantation and in chemotherapy. MDF has mol.wt.
 CC 55-57KD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
 CC contains an amino acid sequence comprising at least one of the
 CC sequences AAR57107-R57115.
 CC
 SQ Sequence 10 AA;
 XX
 QY Query Match 88.0%; Score 44; DB 15; Length 10;
 Best Local Similarity 90.0%; Pred. No. 0.022;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 AVAMMHQXK 10
 DB 1 AVAMMHQXK 10

RESULT 6
 AAY08255
 ID AAY08255 standard; Protein: 380 AA.
 XX
 AC AAY08255;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Rat megasin protein.
 XX
 KW Megasin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 KM human; rat; murine.
 XX
 OS Rattus rattus.
 XX
 PN WO9915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98WO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 1999-276983/23.

DR N-PSDB; AAX56712.
 XX
 PT Mesgin protein expressed specifically in mesangial cells
 XX
 PS Claim 1; Page 69-72; 100pp; Japanese.
 XX
 CC This invention describes the isolation of novel mesgin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IGA nephropathy.
 XX
 SQ Sequence 380 AA;
 QY
 Db 1 AVAMMHQERK 10
 204 AVNMHQR 213
 Query Match 82.0%; Score 41; DB 20; Length 380;
 Best Local Similarity 80.0%; Pred. No. 3.8;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 7
 AAB24150
 ID AAB24150 standard; Protein: 380 AA.
 AC AAB24150;
 DT 30-JAN-2001 (first entry)
 DE Rat mesgin protein sequence SEQ ID NO:19.
 KW Mesgin; mesangium-predominant gene; serpin regulated; nephropathy;
 KM IGA; immunoglobulin A; detection; renal function; renal disorder;
 XX diagnosis; biological sample; blood; urine.
 OS Rattus norvegicus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 51 /note- "unspecified"
 FT Misc-difference 94 /note- "unspecified"
 FT
 PN WO200057189-A1.
 PD 28-SEP-2000.
 XX
 PF 17-MAR-2000; 2000WO-JP01646.
 XX
 PR 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KUROO) KUROKAWA K.
 PA (FUSO) FUSO PHARM IND LTD.
 PA (MIYA) MIYATA T.
 PI Miyata T;
 PI
 DR WPI: 2000-611642/58.
 DR N-PSDB; AAC55238.
 XX
 PT Evaluating renal function comprises assaying mesgin protein in
 PT biological sample
 XX
 PS Example 2; Page 81-84; 93pp; Japanese.
 XX
 CC The present invention describes a method for evaluating renal function.
 CC The method comprises assaying mesgin protein in biological sample. Also
 CC described are: (1) use of an anti-mesgin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting mesgin protein comprising:
 CC (a) anti-mesgin protein antibody attached to solid magnetic particles;

CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying mesgin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the rat mesgin
 CC protein, which is given in the exemplification of the present invention.
 XX
 SQ Sequence 380 AA;
 QY
 Db 1 AVAMMHQERK 10
 204 AVNMHQR 213
 Query Match 82.0%; Score 41; DB 21; Length 380;
 Best Local Similarity 80.0%; Pred. No. 3.8;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 8
 AAG64286
 ID AAG64286 standard; Protein: 380 AA.
 AC AAG64286;
 DT 21-SEP-2001 (first entry)
 DE Rat mesgin protein.
 DE
 KW Rat; mesgin; renal mesangial cell; mesangium proliferative nephritis.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200148019-A1.
 PD 05-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-JP09251.
 XX
 PR 28-DEC-1999; 99JP-0373677.
 XX
 PA (KUROO) KUROKAWA K.
 PA (MIYA) MIYATA T.
 PI Miyata T;
 PI
 DR WPI: 2001-425651/45.
 DR N-PSDB; AAH48181.
 XX
 PT New antibody recognizing a partial sequence of rat mesgin protein for
 PT diagnosis of mesangium proliferative nephritis
 XX
 PS Disclosure; Page 54-56; 63pp; Japanese.
 XX
 CC The present invention relates to a novel antibody which recognises a
 CC peptide consisting of residues 341-354 of rat mesgin protein. The present
 CC sequence is the protein sequence for rat mesgin, which was used in the
 CC present invention. Mesgin is highly expressed in renal mesangial cells
 CC and its level is elevated in mesangium proliferative nephritis. Assay of
 CC the serum or urine level using the antibody is therefore indicative of
 CC this type of disorder.
 XX
 SQ Sequence 380 AA;
 QY
 Db 1 AVAMMHQERK 10
 204 AVNMHQR 213
 Query Match 82.0%; Score 41; DB 22; Length 380;
 Best Local Similarity 80.0%; Pred. No. 3.8;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 RESULT 9

AAB83076	ID	AAB83076 standard; Protein: 380 AA.
AC	XX	
AA83076:	XX	
DT	XX	
10-JUL-2001	XX	(first entry)
DE	XX	Rat mesgln protein.
KW	XX	Rat: mesgln; mesangial cell proliferative nephritis; nephrotropic;
OS	XX	transgenic mouse; glomerular disease; animal model; drug screening.
WO200124628-A1.	XX	
12-APR-2001.	XX	
06-OCT-2000; 2000WO-JP06988.	XX	
06-OCT-1999; 99JP-0285736.	XX	
(KUROO/) KUROKAWA K.	XX	
(MIYATA) MIYATA T.	XX	
Miyata T;	XX	
WPJ: 2001-300136/31.	XX	
N-PSDB: AAF82439.	XX	
Mouse model for mesangial cell proliferative nephritis for development	PT	
and screening of new treatments -	PT	
Disclosure: Page 48-50; 62pp; Japanese.	PS	
The present sequence is rat mesgln. The human mesgln coding	CC	
sequence may be introduced into a mouse to produce an animal model of	CC	
mesangial cell proliferative nephritis. The symptoms include	CC	
enlargement of the mesangial base region, sedimentation of an immune	CC	
complex and an increase in mesangial cells. The animal model is useful	CC	
for analysing the pathology of chronic glomerular diseases and for	CC	
screening compositions for prevention and treatment of the diseases.	CC	
Highly uniform models can be made easily and in large numbers using	CC	
this method.	CC	
Sequence 380 AA;	SO	
Query Match 82.0%; Score 41; DB 22; Length 380;		
Best Local Similarity 80.0%; Pred. No. 3.8;		
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0.		
OY 1 AVAMHGERK 10		
Db 204 AVNMHGERR 213		
RESULT 10		
ABP34239	ID	ABP34239 standard; Protein: 131 AA.
AC	XX	
ABP34239:	XX	
DT	XX	
08-JUL-2002 (first entry)	XX	
Human glycoprotein-like ORF3212 protein, SEQ ID NO:6424.	XX	
Human; ORF: open reading frame; ORFX: drug screening; diagnosis;	KM	
disease monitoring; cytokine; cell proliferation; cell differentiation;	KM	
immune modulation; haematopoiesis regulation; tissue growth;	KM	
angiogenesis; activin; inhibin; chemotactic; chemokine; hemostatic;	KM	
thrombolytic; tumour inhibition; bodily characteristic; fertility;	KM	
behaviour; cancer; proliferative disorder; neurological disorder;	KM	
cardiovascular disease; immune system disorder; organ transplantation;	KM	

KW	tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW	hypothyroidism; cholesterol ester storage disease; infection; vulnerability;
KW	vaccinoidism; cholesteric; antidiabetic; cytostatic; nocitropic;
KW	neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KM	candidant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
KM	dematological; analgesic; vitrucide; antibacterial; fungicide.
XX	
OS	Homo sapiens.
XN	
PN	WO200190366-A2.
XX	
PD	29-NOV-2001.
XX	
PF	24-MAY-2001; 2001WO-US17076.
XX	
PR	24-MAY-2000; 2000US-206690P.
XX	
PA	(CURA-) CURAGEN CORP.
PI	Leach MD, Shinkets RA:
XX	
DR	WPI; 2002-106200/14.
DR	N-PSDB; ABN78265.
XX	
PT	Novel human polypeptides and polynucleotides useful for diagnosing,
PT	preventing and treating cardiovascular disease, neurodegenerative,
PT	hyperproliferative disorders and disorders related to organ
PT	transplantation -
XX	
PS	Claim 10; Page 1854; 2508pp; English.
XX	
CC	Sequences ABP31028-ABP35561 represent 4334 novel human proteins
CC	designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC	ABN79587 represent cDNAs encoding them. The invention also encompasses
CC	polypeptides at least 80% identical to the ORF-ORF4534 (collectively
CC	referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC	the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC	polynucleotides, the recombinant production of ORFX proteins, antibodies
CC	specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC	polypeptides, methods of screening for modulators of ORFX expression or
CC	activity, and methods of screening individuals for a predisposition to an
CC	ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC	range of biological activities, such as cytokine, cell proliferation,
CC	cell differentiation, immune modulation, haematopoiesis regulation,
CC	tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC	chemokinetic activity, haemostatic activity, thrombolytic activity,
CC	receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC	and antifibrotic activity, and may also be involved in the determination
CC	of bodily characteristics, fertility and behaviour. ORFX proteins,
CC	nucleic acids and antibodies may be used in the treatment of cancers,
CC	other proliferative disorders such as psoriasis and benign tumours,
CC	neurological disorders such as epilepsy and Alzheimer's disease,
CC	cardiovascular diseases, immune system disorders, disorders related to
CC	organ transplantation, disorders of tissue growth and regeneration,
CC	diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC	storage disease, and infectious diseases caused by viral, bacterial,
CC	fungal and other pathogens. ORFX nucleic acids may also be used as a
CC	source of primers and probes, in the detection of ORFX genomic sequences
CC	or transcripts, in the identification and cloning of homologous
CC	sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC	nucleic acids may additionally be used to produce transgenic animals
CC	which may be useful for studying the function and/or activity of ORFX
CC	protein, and in drug screening. The ORFX proteins may also be used as
CC	immunogens to generate specific antibodies, which are useful in the
CC	diagnosis, treatment and monitoring of ORFX-associated diseases.
XX	
SQ	Sequence 131 AA:
	Query Match 80.0%; Score 40; DB 23; Length 131;
	Best Local Similarity 88.9%; Pred. No. 2;
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 AVAMMHQER 9

Db 90 AVAMHOSR 98

RESULT 11

ID AAY08256 standard; Protein; 368 AA.

AC AAY08256;

DT 14-JUL-1999 (first entry)

DE Mouse megalin protein.

KW Megalin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
human; rat; murine.

OS Mus musculus.

FT Key Location/Qualifiers
1..368
Protein /note="partial sequence"

PN WO915652-A1.

PD 01-APR-1999.

PF 22-SEP-1998; 98WO-JP04269.

PR 22-SEP-1997; 97JP-0275302.

PA (KURO/) KUROKAWA K.
(MIYA/) MIYATA T.

PI Miyata T;

DR WPI; 1999-276983/23.

DR N-PSDB; AAX56714.

PT Megalin protein expressed specifically in mesangial cells

PS Claim 1; Page 76-79; 100pp; Japanese.

CC This invention describes the isolation of novel megalin nucleic acid and
proteins from human, rat and mouse tissue. This protein is expressed
specifically in mesangial cells. The products of the invention are
useful for the treatment and diagnosis of diseases involving mesangial
cells, such as Iga nephropathy.

SQ Sequence 368 AA;

Query Match 74.0%; Score 37; DB 20; Length 368;

Best Local Similarity 77.8%; Pred. No. 23; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10

DB 193 VMMHOSR 201

RESULT 12

ID AAB24151 standard; Protein; 368 AA.

AC AAB24151;

DT 30-JAN-2001 (first entry)

DE Mouse megalin protein sequence SEQ. ID NO:21.

KW Megalin; mesangium-predominant gene; serpin regulated; nephropathy;
Iga; immunoglobulin A; detection; renal function; renal disorder;
diagnosis; biological sample; blood; urine.

XX Mus musculus.

PN WO200057189-A1.

PD 28-SEP-2000.

PF 17-MAR-2000; 2000WO-JP01646.

PR 19-MAR-1999; 98JP-0075305.

PR 28-OCT-1999; 99JP-0306623.

PA (KURO/) KUROKAWA K.
(FUSO) FUSO PHARM IND. LTD.

PA (MIYA/) MIYATA T.

PI Miyata T;

DR WPI; 2000-61642/58.

DR N-PSDB; AAC55239.

PT Evaluating renal function comprises assaying megalin protein in
biological sample

PS Disclosure; Page 89-91; 93pp; Japanese.

CC The present invention describes a method for evaluating renal function.
The method comprises assaying megalin protein in biological sample. Also
described are: (1) use of a anti-megalin protein antibody for diagnosing
renal function; and (2) a kit for detecting megalin protein comprising:
(a) anti-megalin protein antibody attached to solid magnetic particles;
(b) direct or indirect fixing for the antibody to the particles; and
(c) a magnet. The process is useful for evaluating renal function and
diagnosing renal disorders by assaying megalin protein in biological
samples (preferably urine or blood). The process is reproducible and
gives accurate results. The present sequence represents the mouse megalin
protein, which is given in the exemplification of the present invention.

SQ Sequence 368 AA;

Query Match 74.0%; Score 37; DB 21; Length 368;

Best Local Similarity 77.8%; Pred. No. 23; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10

DB 193 VMMHOSR 201

RESULT 13

ID AAB3077 standard; Protein; 368 AA.

AC AAB3077;

DT 10-JUL-2001 (first entry)

DE Murine megalin protein.

KW Mouse; megalin; mesangial cell proliferative nephritis; nephrotic;
transgenic mouse; glomerular disease; animal model; drug screening.

OS Mus musculus.

PN WO200124628-A1.

PD 12-APR-2001.

PF 06-OCT-2000; 2000WO-JP06988.

PR 06-OCT-1999; 99JP-0285736.

PA (KURO/) KUROKAWA K.

PA (MIVA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2001-300136/31.
 DR N-PSDB: AAF82440.
 XX
 PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments
 XX
 PS Disclosure: Page 52-53; 62pp: Japanese.
 XX
 CC The present sequence is murine mesangin protein. The human mesangin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 XX
 SQ Sequence 368 AA;
 Query Match 74.0%; Score 37; DB 22; Length 368;
 Best Local Similarity 77.8%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 2 VAMMHOERK 10
 1 11111111:
 Db 193 VNMHMOERR 201

RESULT 14
 ABG11324
 ID ABG11324 standard; Protein: 118 AA.
 XX
 AC ABG11324;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #11315.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEO INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS75511.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 41683; 103pp: English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 118 AA;
 Query Match 72.0%; Score 36; DB 22; Length 118;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AVAMMHOERK 10
 1 11111111:
 Db 49 AIAVLHERR 58

RESULT 15
 AAY11676
 ID AAY11676 standard; Protein: 67 AA.
 XX
 AC AAY11676;
 XX
 DT 16-JUN-1999 (first entry)
 XX
 DE Human 5' EST secreted protein SEQ ID NO:328.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haemopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; hemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX
 OS Homo sapiens.
 XX
 PN WO9906439-A2.
 XX
 PD 11-FEB-1999.
 XX
 PF 31-JUL-1998; 98WO-IB01233.
 XX
 PR 01-AUG-1997; 97US-0904468.
 XX
 PA (GEST) GENSET.
 XX
 PI Duclert A, Dumas Milne Edwards J, Lacroix B;
 XX
 DR WPI: 1999-153700/13.
 DR N-PSDB: AAX40394.
 XX
 PT New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from liver, lung, large intestine, colon,
 PT thyroid and pancreas tissue
 XX
 PS Claim 27; Page 395-396; 398pp: English.
 XX
 CC AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY11533 to

CC AAY11679, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, hematopoiesis regulating
 CC activity, tissue growth regulation activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX

SQ Sequence 67 AA;

Query Match 68.08; Score 34; DB 20; Length 67;
 Best Local Similarity 55.68; Pred. No. 16;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMMHOER 9
 1:1::1:1
 Db 18 AINVLHEER 26

Search completed: July 11, 2003, 11:54:26
 Job time : 6.73805 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 1.78794 Seconds
(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-9

Perfect score: 50

Sequence: 1 AVAMHQRK 10

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5
262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	10	2	US-08-611-977-9
2	50	100.0	380	2	US-08-472-659-34
3	50	100.0	380	2	US-08-474-661-34
4	50	100.0	380	2	US-08-611-977-34
5	44	88.0	10	2	US-08-472-659-9
6	44	88.0	10	2	US-08-474-661-9
7	33	66.0	253	2	US-08-685-992-10
8	33	66.0	253	2	US-09-144-925-10
9	33	66.0	793	1	US-08-015-985-3
10	32	64.0	30	1	US-08-287-959-14
11	32	64.0	235	3	US-08-807-151-3
12	32	64.0	235	4	US-09-478-957-3
13	32	64.0	236	1	US-08-015-985-6
14	32	64.0	254	2	US-08-685-992-9
15	32	64.0	254	2	US-09-144-925-9
16	32	64.0	299	2	US-08-773-608A-2
17	32	64.0	798	1	US-08-200-900A-2
18	32	64.0	798	5	PCT-US94-00616-2
19	32	64.0	802	1	US-08-015-985-1
20	32	64.0	1657	1	US-08-287-959-1
21	32	64.0	2071	4	US-09-415-522-6
22	31	62.0	673	4	US-09-196-387-8
23	31	62.0	949	4	US-09-196-387-10
24	31	62.0	1327	4	US-09-196-387-2
25	30	60.0	235	4	US-08-944-483-65
26	29	58.0	166	1	US-08-339-152A-24
27	29	58.0	166	2	US-08-007-999B-13

ALIGNMENTS

28	29	58.0	166	2	US-08-689-276A-13	Sequence 13, Appl
29	29	58.0	171	4	US-09-724-864-59	Sequence 59, Appl
30	29	58.0	238	1	US-08-289-999A-2	Sequence 2, Appl1
31	29	58.0	238	1	US-08-684-203-2	Sequence 2, Appl1
32	29	58.0	238	2	US-08-878-283-2	Sequence 2, Appl1
33	29	58.0	238	3	US-09-082-021-2	Sequence 2, Appl1
34	29	58.0	238	4	US-09-182-616-2	Sequence 2, Appl1
35	29	58.0	245	2	US-08-825-781-1	Sequence 1, Appl1
36	29	58.0	281	4	US-08-961-083-196	Sequence 166, App
37	29	58.0	511	2	US-09-073-362-1	Sequence 1, Appl1
38	29	58.0	511	2	US-09-243-920-1	Sequence 1, Appl1
39	29	58.0	539	3	US-09-057-969-4	Sequence 4, Appl1
40	29	58.0	587	2	US-08-436-664-23	Sequence 23, Appl
41	29	58.0	587	3	US-09-135-642-23	Sequence 23, Appl
42	29	58.0	587	3	US-08-394-232A-23	Sequence 23, Appl
43	29	58.0	587	5	PCT-US95-04080-23	Sequence 23, Appl
44	29	58.0	588	2	US-08-436-664-25	Sequence 25, Appl
45	29	58.0	588	3	US-09-135-642-25	Sequence 25, Appl

RESULT 1
US-08-611-977-9
Sequence 9, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uno
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
State: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611, 977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091, 028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-9

Query Match 100.0%; Score 50; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
DB 1 AVAMMHQERK 10

RESULT 2
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 50; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
DB 204 AVAMMHQERK 213

RESULT 3
US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 50; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
DB 204 AVAMMHQERK 213

RESULT 4

US-08-611-977-34
; Sequence 34, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROJOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhito
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-977-34

Query Match 100.0%; Score 50; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMMHOERK 10
Db 204 AVAMMHOERK 213

RESULT 5
US-08-472-659-9
; Sequence 9, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROJOKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, No. 5831030uhito
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5831030oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-472-659-9

Query Match 88.0%; Score 44; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.0046;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVAMMHOERK 10
Db 1 AVAMMHOERK 10

RESULT 6
US-08-474-661-9
; Sequence 9, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROJOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhito
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5874253oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: George Mason Bldg., Washington & Prince Sts.
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,661
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;;
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 5-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REA, TERESA STANER
;; REGISTRATION NUMBER: 30,427
;; REFERENCE/DOCKET NUMBER: 001560-204
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-6620
;;
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
;; US-08-474-661-9

Query Match 88.0%; Score 44; DB 2; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.0046; 1; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 1;

QY 1 AVAMMHQERR 10
DB 1 AVAMMHQXRR 10

RESULT 7
US-08-685-992-10
; Sequence 10, Application US/08685992
; Patent No. 5912138
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,992

;; FILING DATE: 25-JUL-1996
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: CSHL96-03
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 781-861-6240
;; TELEFAX: 781-861-9540
;;
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 253 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
;; US-08-685-992-10

Query Match 66.0%; Score 33; DB 2; Length 253;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAMMHQERR 10
DB 203 VAMMHQERR 211

RESULT 8
US-09-144-925-10
; Sequence 10, Application US/09144925
; Patent No. 5951979
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas
; APPLICANT: Flint, Andrew J.
; TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02421-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,925
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/685,992
; FILING DATE: JULY 25, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-144-925-10

Query Match 66.0%; Score 33; DB 2; Length 23;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMHORK 10
DB 203 MAMHARK 211

RESULT 9
US-08-015-985-3
Sequence 3, Application US/08015985
Patent No. 553886
GENERAL INFORMATION:
APPLICANT: Schlesinger, Joseph
APPLICANT: Sap, Jan M.
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/015,985
FILING DATE: 10-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 793 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-015-985-3

Query Match 66.0%; Score 33; DB 1; Length 793;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MMHORK 10
DB 452 MMHARK 458

RESULT 10
US-08-287-959-14
Sequence 14, Application US/08287959
Patent No. 5639651
GENERAL INFORMATION:
APPLICANT: Weisbach, Lawrence
APPLICANT: Bernards, Andre
APPLICANT: Settleman, Jeffrey
TITLE OF INVENTION: GAP-RELATED GENE

NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/287,959
FILING DATE: August 9, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/181001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-287-959-14

Query Match 64.0%; Score 32; DB 1; Length 30;
Best Local Similarity 60.0%; Pred. No. 3.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVAMHORK 10
DB 10 SLARHORK 19

RESULT 11
US-08-807-151-3
Sequence 3, Application US/08807151
Patent No. 604303
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL HUMAN PROSTATE-ASSOCIATED
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/807,151
FILING DATE: Filed Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0227 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 416132
US-08-807-151-3

Query Match 64.0%; Score 32; DB 3; Length 235;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VAMHGERK 10
:|||||
Db 93 IAMMHLEMK 101

RESULT 12
US-09-478-957-3
Sequence 3, Application US/09478957
Patent No. 6350448
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL HUMAN PROSTATE-ASSOCIATED
TITLE OF INVENTION: PROTEASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/478,957
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/807,151
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0227 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 416132
US-09-478-957-3

Query Match 64.0%; Score 32; DB 4; Length 235;

Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VAMHGERK 10
:|||||
Db 93 IAMMHLEMK 101

RESULT 13
US-08-015-985-6
Sequence 6, Application US/08015985
Patent No. 553886
GENERAL INFORMATION:
APPLICANT: Schlusser, Joseph
APPLICANT: Sap, Jan M.
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
TITLE OF INVENTION: PHOSPHATASE-ALPHA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/015,985
FILING DATE: 10-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 780-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-015-985-6

Query Match 64.0%; Score 32; DB 1; Length 236;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MMHGERK 10
:|||||
Db 197 MMHTEKK 203

RESULT 14
US-08-685-992-9
Sequence 9, Application US/08685992
Patent No. 5912138
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas
APPLICANT: Flint, Andrew J.
TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millitia Drive

CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,992
FILING DATE: 25-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: CSHL96-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-685-992-9

Query Match 64.0%; Score 32; DB 2; Length 254;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MMHQERK 10
|||
Db 206 MMHTERK 212

RESULT 15
US-09-144-925-9
Sequence 9, Application US/09144925
Patent No. 5951979
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas
APPLICANT: Flint, Andrew J.
TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millita Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02421-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,925
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/685,992
FILING DATE: July 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: CSHL96-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-144-925-9

Query Match 64.0%; Score 32; DB 2; Length 254;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MMHQERK 10
|||
Db 206 MMHTERK 212

Search completed: July 11, 2003, 12:02:12
Job time: 2.78794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 2.9319 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-9
Perfect score: 50
Sequence: 1 AVAMMHQERK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications-AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	50	100.0	10	9	US-10-091-442-9
2	50	100.0	10	9	US-09-140-719-9
3	50	100.0	380	9	US-10-091-442-34
4	50	100.0	380	10	US-09-140-719-34
5	33	66.0	309	10	US-09-788-626-11
6	32	64.0	235	9	US-09-988-975A-6
7	32	64.0	279	9	US-09-858-332-15
8	32	64.0	310	10	US-09-788-626-10
9	31	62.0	318	10	US-09-919-497-91
10	31	62.0	318	10	US-09-919-497-92
11	31	62.0	370	9	US-10-128-714-8591
12	31	62.0	673	10	US-09-841-835-8
13	31	62.0	718	10	US-09-815-242-10901
14	31	62.0	949	10	US-09-841-835-10
15	31	62.0	1347	9	US-09-972-115A-8
16	31	62.0	1347	10	US-09-841-835-2
17	31	62.0	54	10	US-09-864-761-47786
18	30	60.0	82	9	US-09-956-622A-38
19	30	60.0			Sequence 38, Appl

20	30	60.0	377	9	US-10-165-605A-27	Sequence 27, Appl
21	30	60.0	377	10	US-09-910-430-27	Sequence 27, Appl
22	30	60.0	427	8	US-08-808-031A-40	Sequence 40, Appl
23	30	60.0	427	10	US-09-925-297-684	Sequence 684, App
24	30	60.0	1005	9	US-09-934-070-8	Sequence 8, Appl
25	30	60.0	1019	9	US-10-157-031-267	Sequence 267, App
26	30	60.0	1019	9	US-09-776-191-64	Sequence 64, Appl
27	29	58.0	155	9	US-09-866-050A-699	Sequence 699, App
28	29	58.0	173	10	US-09-867-550-888	Sequence 888, App
29	29	58.0	261	9	US-10-156-761-8516	Sequence 8516, App
30	29	58.0	281	10	US-09-765-272-196	Sequence 196, App
31	29	58.0	285	9	US-10-156-761-12080	Sequence 12080, A
32	29	58.0	314	9	US-09-925-299-934	Sequence 934, App
33	29	58.0	314	9	US-10-106-698-4725	Sequence 4725, App
34	29	58.0	314	10	US-09-925-299-934	Sequence 934, App
35	29	58.0	332	10	US-09-764-853-606	Sequence 606, App
36	29	58.0	356	9	US-10-123-965B-5	Sequence 5, Appl
37	29	58.0	427	10	US-09-815-242-13443	Sequence 13443, A
38	29	58.0	427	10	US-09-815-242-13640	Sequence 13640, A
39	29	58.0	458	10	US-09-815-242-11759	Sequence 11759, A
40	29	58.0	511	9	US-10-163-866-41	Sequence 41, Appl
41	29	58.0	511	9	US-10-163-866-42	Sequence 42, Appl
42	29	58.0	638	9	US-10-101-464A-74	Sequence 74, Appl
43	29	58.0	734	9	US-10-040-862-10463	Sequence 10463, A
44	29	58.0	773	9	US-10-112-286-2	Sequence 2, Appl
45	28	56.0	50	9	US-10-106-698-7062	Sequence 7062, App

ALIGNMENTS

RESULT 1
US-10-091-442-9
Sequence 9, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Eiyuki
TSUKUDOKA, No. US20020164711A1uo
NAKAZATO, Hiroshi
MURA, Kenju
ISHIDA, No. US20020164711A1uhiro
KURIHARA, Tatsuya
YAMACHI, Kozo
YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-091-442-9

Query Match 100.0%; Score 50; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
DB 1 AVAMMHQERK 10

RESULT 2
US-09-140-719-9
Sequence 9, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUROOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-9

Query Match 100.0%; Score 50; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
DB 1 AVAMMHQERK 10

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUROOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uh1ro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-MAR-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 50; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMHGERK 10
DB 204 AVAMHGERK 213

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUROOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1om1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 50; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMHGERK 10
DB 204 AVAMHGERK 213

RESULT 5
US-09-788-626-11
Sequence 11, Application US/09788626
Patent No. US20020009762A1
GENERAL INFORMATION:
APPLICANT: Flint, Andrew J.
APPLICANT: COOL, Deborah E.
TITLE OF INVENTION: IMPROVED ASSAY FOR PROTEIN TYROSINE
PHOSPHATASES
FILE REFERENCE: 200125,401
CURRENT APPLICATION NUMBER: US/09/788,626
CURRENT FILING DATE: 2001-02-13
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 309
TYPE: PRT
ORGANISM: Homo sapiens
US-09-788-626-11

Query Match 66.0%; Score 33; DB 10; Length 309;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMHGERK 10
DB 203 VAMHGERK 211

RESULT 6
US-09-988-975A-6
Sequence 6, Application US/09988975A
Patent No. US20020119531A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti G.
TITLE OF INVENTION: PROSTATE-ASSOCIATED PROTEASE ANTIBODY
FILE REFERENCE: PP-0227-2 CIP
CURRENT APPLICATION NUMBER: US/09/988,975A
CURRENT FILING DATE: 2001-11-19
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Genbank ID No. US20020119531A1 g416132
US-09-988-975A-6

Query Match 64.0%; Score 32; DB 10; Length 235;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 VAMHGERK 10
DB 93 VAMHGERK 101

RESULT 7

US-09-858-332-15
; Sequence 15, Application US/09858332
; Patent No. US20020164718A1
; GENERAL INFORMATION:
; APPLICANT: Tchaqa, Grigory S.
; APPLICANT: Tchaqa, Grigory S.
; TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for
; FILE REFERENCE: CLON056CIP
; CURRENT APPLICATION NUMBER: US/09/858,332
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 09/404,017
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,867
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-858-332-15

Query Match 64.0%; Score 32; DB 9; Length 279;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10
Db 122 IMMHELEK 130

RESULT 8
US-09-788-626-10
; Sequence 10, Application US/09788626
; Patent No. US2002009762A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Andrew J.
; APPLICANT: Cool, Deborah E.
; TITLE OF INVENTION: IMPROVED ASSAY FOR PROTEIN TYROSINE
; FILE REFERENCE: 200125,401
; CURRENT APPLICATION NUMBER: US/09/788,626
; CURRENT FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-626-10

Query Match 64.0%; Score 32; DB 10; Length 310;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MMHOERK 10
Db 206 MMHTEK 212

RESULT 9
US-09-919-497-91
; Sequence 91, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-91

Query Match 62.0%; Score 31; DB 10; Length 318;
Best Local Similarity 62.5%; Pred. No. 14e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 AMMHOERK 10
Db 190 ALIHKEK 197

RESULT 10
US-09-919-497-92
; Sequence 92, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-92

Query Match 62.0%; Score 31; DB 10; Length 318;
Best Local Similarity 62.5%; Pred. No. 14e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 AMMHOERK 10
Db 190 ALIHKEK 197

RESULT 11
US-10-128-714-8591
; Sequence 8591, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Weng
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362

;; PRIOR FILING DATE: 2001-08-31
;; NUMBER OF SEQ ID NOS: 8603
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 8591
;; LENGTH: 370
;; TYPE: PR
;; ORGANISM: Aspergillus fumigatus
US-10-128-714-8591

Query Match 62.0%; Score 31; DB 9; Length 370;
Best Local Similarity 75.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMOER 9
DB 311 VAMRHKR 318

RESULT 12
US-10-128-714-3591
;; Sequence 3591, Application US/10128714
;; Publication No. US20030119013A1
;; GENERAL INFORMATION:
;; APPLICANT: Jiang, Bo
;; APPLICANT: Hu, Meng
;; APPLICANT: Tishkoff, Daniel
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Erosimkin, Alexey M
;; APPLICANT: Lemieux, Sebastien M
;; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;; FILE REFERENCE: 10182-018-999
;; CURRENT APPLICATION NUMBER: US/10/128,714
;; CURRENT FILING DATE: 2002-04-23
;; PRIOR APPLICATION NUMBER: US 60/285,697
;; PRIOR FILING DATE: 2001-04-23
;; PRIOR APPLICATION NUMBER: US 60/287,066
;; PRIOR FILING DATE: 2001-04-27
;; PRIOR APPLICATION NUMBER: US 60/295,890
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: US 60/303,899
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: US 60/316,362
;; PRIOR FILING DATE: 2001-08-31
;; NUMBER OF SEQ ID NOS: 8603
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 3591
;; LENGTH: 445
;; TYPE: PR
;; ORGANISM: Aspergillus fumigatus
US-10-128-714-3591

Query Match 62.0%; Score 31; DB 9; Length 445;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMOER 9
DB 386 VAMRHKR 393

RESULT 13
US-09-841-835-8
;; Sequence 8, Application US/09841835
;; Patent No. US20020076795A1
;; GENERAL INFORMATION:
;; APPLICANT: de Lange, Titia
;; APPLICANT: Smith, Susan
;; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Klauber & Jackson

;; STREET: 411 Hackensack Avenue, 4th Floor
;; CITY: Hackensack
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07601
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/841,835
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/196,387
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Jackson Esq., David A.
;; REGISTRATION NUMBER: 26,742
;; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-487-5800
;; TELEFAX: 201-343-1684
;; TELEX: 133521
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 673 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-841-835-8

Query Match 62.0%; Score 31; DB 10; Length 673;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVAMMOER 10
DB 529 AVASLHPRK 538

RESULT 14
US-09-815-242-10901
;; Sequence 10901, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.01A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 1410
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 10901
 LENGTH: 718
 TYPE: PRT
 ORGANISM: Enterococcus faecalis
 US-09-815-242-10901

Job time : 3.93139 secs

Query Match 62.0%; Score 31; DB 10; Length 718;
 Best Local Similarity 66.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVAMMOER 9
 122 AVAMMOER 130

RESULT 15
 US-09-841-835-10
 Sequence 10, Application US/09841835
 Patent No. US20020076795A1
 GENERAL INFORMATION:
 APPLICANT: de Lange, Titia
 APPLICANT: Smith, Susan
 TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
 TITLE OF INVENTION: OF USE THEREOF
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 411 Hackensack Avenue, 4th Floor
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/841,835
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/196,387
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 949 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-841-835-10

Query Match 62.0%; Score 31; DB 10; Length 949;
 Best Local Similarity 60.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVAMMOER 10
 529 AVASLHPRK 538

Search completed: July 11, 2003, 12:37:35

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 2.01663 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-9
Perfect score: 50
Sequence: 1 AVAMMHQERK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	70.0	374	2 T34435	hypothetical prote
2	35	70.0	578	2 H86177	hypothetical prote
3	35	70.0	602	2 A84744	hypothetical prote
4	35	70.0	1561	2 T00248	zinc finger protei
5	33	66.0	275	2 H70234	replicative DNA he
6	33	66.0	281	2 AD2052	hypothetical prote
7	33	66.0	327	1 E35270	6-phosphotrioxin
8	33	66.0	337	1 D72084	holiday junction
9	33	66.0	337	2 D86539	holiday junction
10	33	66.0	337	2 A81586	holiday junction
11	33	66.0	395	2 PC1143	finger protein zfc
12	33	66.0	700	1 S12053	protein-tyrosine-p
13	33	66.0	796	1 JC1285	protein-tyrosine-p
14	33	66.0	829	1 A47373	protein-tyrosine-p
15	32	64.0	229	2 C86399	protein F17L21.16
16	32	64.0	252	2 A13408	glycoprotein/poly
17	32	64.0	298	2 E70912	hypothetical prote
18	32	64.0	384	2 H64612	hypothetical prote
19	32	64.0	401	2 A47014	probable arylsulfa
20	32	64.0	431	1 JX0364	antithrombin III -
21	32	64.0	474	2 A40721	neuroblast prolif
22	32	64.0	486	2 F83390	hypothetical prote
23	32	64.0	496	2 JC5170	probable methylgl
24	32	64.0	656	1 B44954	fumarate reductase
25	32	64.0	802	1 A36065	protein-tyrosine-p
26	32	64.0	1035	1 A43090	enteropeptidase (E
27	32	64.0	1126	1 S49208	transmembrane prot
28	32	64.0	1657	2 A54854	Ras GTPase activat
29	31	62.0	138	1 TTRB0B	thyrotropin beta c

30	31	62.0	138	1 TTRPB	thyrotropin beta c
31	31	62.0	148	2 T24628	hypothetical prote
32	31	62.0	216	2 AE0648	nitrate/nitrite re
33	31	62.0	217	2 B83944	resolvase BR2354 I
34	31	62.0	217	2 F83502	hypothetical prote
35	31	62.0	236	2 A82287	phosphate transpor
36	31	62.0	253	1 G69109	conserved hypothet
37	31	62.0	301	2 T36520	hypothetical prote
38	31	62.0	318	1 K1HUR2	ribose-phosphate d
39	31	62.0	318	1 K1HUR2	ribose-phosphate d
40	31	62.0	318	1 K1HUR1	ribose-phosphate d
41	31	62.0	318	1 K1HUR1	ribose-phosphate d
42	31	62.0	318	1 K1HUR3	ribose-phosphate d
43	31	62.0	334	2 C69148	drpp-glucose 4,6-d
44	31	62.0	397	2 D64690	signal-transducing
45	31	62.0	506	2 D71461	hypothetical prote

ALIGNMENTS

RESULT 1
T34435
hypothetical protein K11H12.9 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C.Accession: T34435
R.Bradshtaw, H.
submitted to the EMBL Data Library, February 1997
A.Description: The sequence of C. elegans cosmid K11H12.
A.Reference number: 221526
A.Accession: T34435
A>Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-374 <BRA>
A.Cross-references: EMBL:U88168; PIDN:AAC24403.1; GSPDB:GN00022; CESP:K11H12.9
A.Experimental source: strain Bristol N2; clone K11H12
C.Genetics:
A.Gene: CESP:K11H12.9
A.Map position: 4
A.Introns: 65/2; 218/3; 254/1; 301/1; 355/1

Query Match
Best Local Similarity 70.0%; Score 35; DB 2; Length 374;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AVAMMHQ 7
Db 206 AVAMMHQ 212
|||||||

RESULT 2
H86177
hypothetical protein [imported] - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C.Accession: H86177
R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000

A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Malt, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A.Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tello
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A.Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A.Reference number: A86141; MIMD:21016719; PMID:11130712
A.Accession: H86177
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-578 <STO>
A.Cross-references: GB:AE005172; NID:g2494130; PIDN:AA80639.1; GSPDB:GN00141

```

C:Genetics:
A:Map position: 1

Query Match      70.0% Score 35; DB 2; Length 578;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 AVAMMHQERK 10
        ||| :|||
Db       29 AVAMVHSERK 38

RESULT 3
AB4744
hypothetical protein At2g33320 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: AB4744
R:Rilln, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;
Euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: AB4744
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-602 <STO>
A:Cross-references: GB:AE002093; NID:g2459415; PIDN:AAB80650.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g33320
A:Map position: 2

Query Match      70.0% Score 35; DB 2; Length 602;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      1 AVAMMHQERK 10
        ||| :|||
Db       31 AVAMVHSERK 40

RESULT 4
T00248
Zinc finger protein w1zL - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 05-Nov-1999
C:Accession: T00248
R:Matsumoto, K.; Ishii, N.; Yoshida, S.; Shiosaka, S.; Wanaka, A.; Tohyama, M.
Submitted to the EMBL Data Library, March 1998
A>Description: Molecular cloning and distinct developmental expression pattern of splice
A:Reference number: Z14130
A:Accession: T00248
A>Status: preliminary; translated from GR/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1561 <MAT>
A:Cross-references: EMBL:AB012265; NID:d1227740; PIDN:BAA32790.1; PTD:d1033756
C:Genetics:
A:Experimental source: Brain
A:Gene: w1z

Query Match      70.0% Score 35; DB 2; Length 1561;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      1 AVAMMHQERK 10
        |::| |::|||
Db       604 AMALKHERK 613

RESULT 5
I70234
replicative DNA helicase homolog - Lyme disease spirochete plasmid G/lp28-2

C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: H70234
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathgre, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genome sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:9805943; PMID:9403685
A:Accession: H70234
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-275 <RLF>
A:Cross-references: GB:AE000766; NID:g2690008; PIDN:AAC66061.1; PID:g2690018; TIGR:BB
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match      66.0% Score 33; DB 2; Length 275;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      2 VAMMHQERK 10
        | :|||||
Db       223 VFRLHQERK 231

RESULT 6
AD2052
hypothetical protein all1970 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A>Note: Nostoc sp. Strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AD2052
R:Keneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2052
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB73669.1; PID:g17131060; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1970

Query Match      66.0% Score 33; DB 2; Length 281;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      2 VAMMHQERK 10
        || |||||
Db       174 YAKHHQERK 182

RESULT 7
E35270
6-phosphofructokinase (EC 2.7.1.11) - Spiroplasma citri
N:Alternate names: phosphofructokinase I; phosphohexokinase
C:Species: Spiroplasma citri
C>Date: 17-Aug-1990 #sequence_revision 14-Jul-1994 #text_change 07-Dec-1999
C:Accession: E35270
R:Chevalier, C.; Saillard, C.; Bove, J.M.
J. Bacteriol. 172, 2693-2703, 1990
A>Title: Organization and nucleotide sequences of the Spiroplasma citri genes for rib
in.
A:Reference number: A35270; MUID:90236934; PMID:2139649
A:Accession: E35270
A:Molecule type: DNA
A:Residues: 1-327 <CHE>
```

A:Cross-references: GB:M31161; NID:9152884; PID:9152889
A:Note: the authors translated the codon GCT for residue 195 as Ile
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: 6-phosphofructokinase; 6-phosphofructokinase 1 homology
C:Keywords: ATP; phosphotransferase
F:5-280/Domain: 6-phosphofructokinase 1 homology <6PFL>

Query Match 66.0%; Score 33; DB 1; Length 327;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10
Db 206 VAMLHODKK 214

RESULT 8
D72084
holliday junction helicase - Chlamydia pneumoniae (strain CML029)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: D72084
R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72084
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <RNA>
A:Cross-references: GB:AE001623; GB:AE001363; NID:94376662; PIDN:AMD18533.1; PID:9437666
A:Experimental source: strain CML029
C:Genetics:
A:Gene: ruvB
C:Superfamily: ruvB protein

Query Match 66.0%; Score 33; DB 2; Length 337;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10
Db 5 VAVLHODKK 13

RESULT 9

D86539
holliday junction helicase [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D86539
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; IS
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: D86539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <STO>
A:Cross-references: GB:BA000008; NID:98978762; PIDN:BA98598.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: ruvB
C:Superfamily: ruvB protein

Query Match 66.0%; Score 33; DB 2; Length 337;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10
Db 5 VAVLHODKK 13

RESULT 10

A81586
Holliday junction DNA helicase ruvB CP0365 [imported] - Chlamydia pneumoniae (str
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: A81586
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heldelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Koloney, J.; McClarty, G.; Salzb
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <REA>
A:Cross-references: GB:AE002198; GB:AE002161; NID:97189279; PIDN:AAF38214.1; PID:9718
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0365
C:Superfamily: ruvB protein

Query Match 66.0%; Score 33; DB 2; Length 337;
Best Local Similarity 55.6%; Pred. No. 29;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAMMHOERK 10
Db 5 VAVLHODKK 13

RESULT 11

PC1143
finger protein zfc - American alligator (fragment)
C:Species: Alligator mississippiensis (American alligator)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 01-Dec-2000
C:Accession: PC1143
R:Valley, E.M.A.; Mueller, U.; Ferguson, M.W.J.; Sharpe, P.T.
Gene 119, 221-228, 1992
A:Title: Cloning and expression analysis of two zfy-related zinc finger genes from Al
A:Reference number: PC1143; MUID:93012990; PMID:1398103
A:Accession: PC1143
A:Molecule type: DNA
A:Residues: 1-395 <VAL>
A:Cross-references: EMBL:X61714; NID:962415; PIDN:CAA43863.1; PID:962416
C:Genetics:
A:Gene: zfc
C:Superfamily: zinc finger protein zfp-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger

Query Match 66.0%; Score 33; DB 2; Length 395;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 AMMHOERK 10
Db 214 ALMHQSK 221

RESULT 12

S12053
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type epsilon precursor - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S12053
R:Kruenger, N.X.; Streuli, M.; Saito, H.
EMBO J. 9, 3241-3252, 1990
A:Title: Structural diversity and evolution of human receptor-like protein tyrosine p
A:Reference number: S12049; MUID:91006018; PMID:2170109
A:Accession: S12053
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-700 <KRU>
 A:Cross-references: GB:X54134; NID:935791; PIDN:CAA38069.1; PID:935792
 C:Genetics:
 A:Gene: GDB:PTPR2
 A:Cross-references: GDB:131385; OMIM:600926
 A:Map position: 10q26-10q26
 C:Function:
 A:Description: catalyzes the hydrolysis of peptidyl-phosphotyrosine to release phosphate
 C:Superfamily: protein-tyrosine-phosphatase, receptor type alpha; leukocyte common antigen
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane prot
 F:1-19/Domain: signal sequence #status predicted <Sig>
 F:20-700/Product: protein-tyrosine-phosphatase; receptor type epsilon #status predicted
 F:47-63/Domain: transmembrane #status predicted <TMN>
 F:78-698/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:159-383/Domain: protein-tyrosine-phosphatase homology <PRP1>
 F:335/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:341/Binding site: substrate phosphatase (Arg) #status predicted
 F:630/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:636/Binding site: substrate phosphatase (Arg) #status predicted

Query Match 66.0%; Score 33; DB 1; Length 700;
 Best Local Similarity 66.7%; Pred. No. 63;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMOERK 10
 DB 352 MAMMAEOK 360

RESULT 13
 JCI285
 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type alpha precursor - rat
 M:Alternate names: protein-tyrosine-phosphatase LRP
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999
 C:Accession: JCI285; S23253; JH0450
 R:Moriyama, T.; Fujiwara, Y.; Imai, E.; Takenaka, M.; Kawanishi, S.; Inoue, T.; Noguchi, T.
 Biochem. Biophys. Res. Commun. 188, 34-39, 1992
 A:Title: cDNA cloning of rat LRP, a receptor like protein tyrosine phosphatase, and evid
 A:Reference number: JCI285; MUID:93038682; PMID:117854
 A:Accession: JCI285
 A:Molecule type: mRNA
 A:Residues: 1-796 <MOR>
 A:Experimental source: kidney
 A:Note: The authors translated the codon TCC for residue 788 as Ala
 R:Hashimoto, N.; Zhang, W. R.; Goldstein, B. J.
 Biochem. J. 284, 569-576, 1992
 A:Title: Insulin receptor and epidermal growth factor receptor dephosphorylation by three
 A:Reference number: S23126; MUID:92287069; PMID:1599438
 A:Accession: S23253
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 254-267, 'I', 269-354, 'T', 356-501 <HMS>
 R:Zhang, W. R.; Goldstein, B. J.
 Biochem. Biophys. Res. Commun. 178, 1291-1297, 1991
 A:Title: Identification of skeletal muscle protein-tyrosine phosphatases by amplificatio
 A:Reference number: JH0450; MUID:91337074; PMID:1651716
 A:Accession: JH0450
 A:Molecule type: mRNA
 A:Residues: 324-354, 'T', 356-434 <ZHA>
 A:Experimental source: skeletal muscle, strain Sprague-Dawley
 C:Function:
 A:Description: catalyzes the hydrolysis of peptidyl-phosphotyrosine to release phosphate
 A:Note: Allows recovery from insulin stimulation by dephosphorylating insulin receptor a
 C:Superfamily: protein-tyrosine-phosphatase, receptor type alpha; leukocyte common antigen
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; receptor; transmembrane prot
 F:1-19/Domain: signal sequence #status predicted <Sig>
 F:20-796/Product: protein-tyrosine-phosphatase; receptor type alpha #status predicted <W
 F:20-148/Domain: extracellular #status predicted <EXT>
 F:149-165/Domain: transmembrane #status predicted <TMN>
 F:166-796/Domain: intracellular #status predicted <INT>
 F:178-794/Domain: leukocyte common antigen cytosolic domain homology <LAC>
 F:259-484/Domain: protein-tyrosine-phosphatase homology <PRP1>

F:552-774/Domain: protein-tyrosine-phosphatase homology <PRP2>
 F:436/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:442/Binding site: substrate phosphatase (Arg) #status predicted
 F:726/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:732/Binding site: substrate phosphatase (Arg) #status predicted

Query Match 66.0%; Score 33; DB 1; Length 796;
 Best Local Similarity 85.7%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MMHSEK 10
 DB 455 MMHSEK 461

RESULT 14
 A47373
 protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type alpha-precursor - mouse
 M:Alternate names: Leukocyte common antigen-related protein LRP; PTPY28; receptor pr
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 21-Jan-2000
 C:Accession: A47373; B35501; A35501; P50367; S40286; A36004; D61180
 R:Hong, E. C.; Mullersman, J. E.; Thomas, M. L.
 Genomics 17, 33-38, 1993
 A:Title: Leukocyte common antigen-related phosphatase (LRP) gene structure: conservat
 A:Reference number: A47373; MUID:94010906; PMID:8406469
 A:Accession: A47373
 A:Molecule type: DNA
 A:Residues: 1-829 <MON>
 A:Cross-references: GB:L13607
 A:Note: Sequence extracted from NCBI backbone (NCBIN:137623, NCBI:137624)
 R:Matthews, R. J.; Cahill, E. D.; Thomas, M. L.
 Proc. Natl. Acad. Sci. U.S.A. 87, 4444-4448, 1990
 A:Title: Identification of an additional member of the protein-tyrosine-phosphatase f
 A:Accession: B35501
 A:Reference number: A35501; MUID:90280391; PMID:2162042
 A:Molecule type: mRNA
 A:Residues: 1-829 <MA1>
 A:Cross-references: GB:M36033; NID:9198876; PIDN:AAA39448.1; PID:9198877; GB:M33671
 A:Note: The authors translated the codon GAT for residue 30 as Tyr
 A:Accession: A35501
 A:Molecule type: mRNA
 A:Residues: 1-267, 'Y', 305-829 <MA2>
 A:Cross-references: GB:M36033; GB:M33671
 R:den Hertog, J.; Pals, C. E. G. M.; Jonk, L. J. C.; Krulder, W.
 Biochem. Biophys. Res. Commun. 184, 1241-1249, 1992
 A:Title: Differential expression of a novel murine non-receptor protein tyrosine phos
 A:Reference number: JH0609; MUID:92272714; PMID:1590786
 A:Accession: P50367
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 322-325, 'G', 327-356 <DEN>
 A:Experimental source: embryonal carcinoma cell, P19 cell
 R:Hendriks, W.; Brugman, C.; Zeeuwen, P.; Schepens, J.; Wieringa, B.
 Submitted to the EMBL Data Library, June 1993
 A:Description: Assessment of the expression levels of murine protein-tyrosine phosphat
 A:Reference number: S40280
 A:Accession: S40285
 A:Molecule type: mRNA
 A:Residues: 358-467 <HEN>
 A:Cross-references: EMBL:223054; NID:9438145; PIDN:CAA80589.1; PID:9438146
 A:Accession: S40286
 A:Molecule type: mRNA
 A:Residues: 651-756 <HE2>
 A:Cross-references: EMBL:223055; NID:9438147; PIDN:CAA80590.1; PID:9438148
 R:Sap, J.; D'Eustachio, P.; Givol, D.; Schlesinger, J.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6112-6116, 1990
 A:Title: Cloning and expression of a widely expressed receptor tyrosine phosphatase.
 A:Reference number: A36004; MUID:90349565; PMID:2166945
 A:Accession: A36004
 A:Molecule type: mRNA
 A:Residues: 1-230, 'L', 232-267, 'Y', 305-410, 'S', 412-829 <SAP>
 A:Cross-references: GB:M34668

Job time : 4.11663 secs

R:YL, T.: Cleveland, J.L.; Inle, J.N.

Blood 78, 2222-2228, 1991

A:Title: Identification of novel protein tyrosine phosphatases of hematopoietic cells by

A:Reference number: A61180; MID:92032882; PMID:1932742

A:Accession: D61180

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 358-467 <YIA>

C:Genetics:

A:Map position: 2

C:Function:

A:Description: catalyzes the hydrolysis of peptidyl-phosphotyrosine to release phosphate

A:Note: allows recovery from insulin stimulation by dephosphorylating insulin receptor a

C:Superfamily: protein-tyrosine-phosphatase, receptor type alpha; leukocyte common antig

C:Keywords: alternative splicing; glycoprotein; phosphoprotein; phosphoric monoester hyd

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-829/Product: protein-tyrosine-phosphatase, receptor type alpha #status predicted <M

F:40-142/Domain: extracellular #status predicted <EXT>

F:143-166/Domain: transmembrane #status predicted <TM>

F:167-829/Domain: intracellular #status predicted <INT>

F:175-827/Domain: leukocyte common antigen cytosolic domain homology <LAC>

F:21,47,51,68,80,86,104,124/Binding site: carbohydrate (asn) (covalent) #status predicte

F:469/Active site: Cys (phosphocysteine intermediate) #status predicted

F:475/Binding site: substrate phosphate (Arg) #status predicted

F:759/Active site: Cys (phosphocysteine intermediate) #status predicted

F:765/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 66.0%; Score 33; DB 1; Length 829;

Best Local Similarity 85.7%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 MMH0ERK 10

DB 488 MMH0ERK 494

RESULT 15

C86399

protein F17L21.16 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C86399

R:Theologos, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maitl, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MID:21016719; PMID:11130712

A:Accession: C86399

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <STO>

A:Cross-references: GB:AEO05172; MID:99802545; PIDN:AAF99747.1; GSPDB:GN00141

C:Genetics:

A:Gene: F17L21.16

A:Map position: 1

Query Match

Best Local Similarity 64.0%; Score 32; DB 2; Length 229;

Best Local Similarity 60.0%; Pred. No. 32;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 AVAMH0ERK 10

DB 24 AVATTH0ERK 33

Search completed: July 11, 2003, 12:00:44

THIS PAGE BLANK (USF...)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 0.956341 Seconds

(Without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-9

Perfect score: 50

Sequence: 1 AVAMHGERK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	50	100.0	SPB7_HUMAN	O75635 homo sapien
2	33	66.0	K6PF_SPICI	P20275 spiroplasma
3	33	66.0	RUVB_CHLPN	Q286F3 chlamydia p
4	33	66.0	PTPE_HUMAN	P23469 homo sapien
5	33	66.0	PTRA_RAT	Q03348 rattus norv
6	33	66.0	PTRA_MOUSE	P18032 mus musculu
7	32	64.0	ICE1_SPOFR	P9116 spodoptera
8	32	64.0	KSF5_ECOLI	Q47334 escherichia
9	32	64.0	CHUR_BACTN	Q02550 bacteroides
10	32	64.0	ANA_DROME	Q26307 drosophila
11	32	64.0	MGLA_TREPA	Q56342 treponema p
12	32	64.0	FRDA_WOLSU	P17412 wolfinella s
13	32	64.0	PTRA_HUMAN	P18433 homo sapien
14	32	64.0	ENTK_BOVIN	P98072 bos taurus
15	32	64.0	HEM_DROME	P51162 drosophila
16	32	64.0	IOGI_HUMAN	P46940 homo sapien
17	32	64.0	IOGI_MOUSE	Q93461 mus musculu
18	31	62.0	TSBH_BOVIN	P11223 bos taurus
19	31	62.0	TSBH_HORSE	Q28336 equus cabal
20	31	62.0	TSBH_PIG	P79357 lama glama
21	31	62.0	TSBH_PIG	P11224 sus scrofa
22	31	62.0	KPR1_HUMAN	P09339 homo sapien
23	31	62.0	KPR2_HUMAN	P11908 homo sapien
24	31	62.0	KPR2_RAT	P09330 rattus norv
25	31	62.0	KPR3_HUMAN	P21108 homo sapien
26	31	62.0	TNKL_HUMAN	O95271 homo sapien
27	31	62.0	MRP6_HUMAN	O95255 homo sapien
28	30	60.0	RNH_ZYMO	O69014 zymononas m
29	30	60.0	YHBU_ECOLI	P45527 escherichia
30	30	60.0	ILVC_CAJGE	Q93461 mus musculu
31	30	60.0	DBH1_MOUSE	P51656 mus musculu
32	30	60.0	TRUA_HUMAN	O93461 mus musculu
33	30	60.0	GALI_STACA	Q93461 mus musculu

ALIGNMENTS

RESULT 1	SPB7_HUMAN	STANDARD	PRT	380 AA.
AC	O75635:			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Megsin (TP55) (Serpin B7).			
GN	SERPINB7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE-97326116; PubMed-9182567;			
RA	Tsujiimoto M., Tsunoka N., Ishida N., Kurihara T., Iwasa F.,			
RA	Yamashiro K., Rogi T., Kodama S., Katsunagi N., Adachi M.,			
RA	Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,			
RA	Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.;			
RT	"Purification, cDNA cloning, and characterization of a new serpin with			
RT	megakaryocyte maturation activity."			
RL	J. Biol. Chem. 272:15373-15380(1997).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
EX	TISUB-Mesangial cells:			
RA	Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,			
RA	Okubo K., Kurokawa K.;			
RT	"A mesangium-predominant gene, megin, is a new serpin upregulated in			
RT	IGA nephropathy."			
RL	J. Clin. Invest. 102:828-836(1998).			
CC	-1- FUNCTION: Might function as an inhibitor of lys-specific			
CC	proteases. Might influence the maturation of megakaryocytes via			
CC	its action as a serpin.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.			
CC	-1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.1sb-sib.ch/announce/			
CC	or send an email to license@1sb-sib.ch).			
CC	-----			
DR	EMBL; D88575; BAA31232.1; -			O93461 mus musculu
DR	EMBL; AF027866; AAC64506.1; -			O93461 mus musculu
DR	HSSP; P05619; 1HE.			P23071 myxococcus
DR	Genew; HGNC:13902; SERPINB7.			P41361 bos taurus
DR	MTM; 603357; -			Q01447 fusarium so
DR	InterPro: IPR000215; Serpin.			O13597 septoria ly
DR	Pfam; PF00079; serpin; 1.			Q62b8 yersinia pe
DR	SMART; SM00093; SERPIN; 1.			P13670 vibrio harv
DR	PROSITE; PS00284; SERPIN; 1.			P98073 homo sapien
DR				P98074 sus scrofa
DR				P97433 mus musculu
DR				P20193 drosophila

KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 347 348 REACTIVE_BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CF605 CRC64;

Query Match 100.0%; Score 50; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMHQRK 10
 DB 204 AVAMHQRK 213

RESULT 2
 ID K6PF.SPICI STANDARD; PRT; 327 AA.
 AC P20275;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 6-phosphofructokinase (EC 2.7.1.11) (Phosphofructokinase)
 DE PFKA.
 GN Spiroplasma citri.
 OS Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
 OC Spiroplasmataceae; Spiroplasma.
 OX NCBI_TaxId=2133;
 RN NCBI
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27556 / R8A2;
 RX MEDLINE=90236934; Pubmed=2139649;
 RA Chevallier C., Sallard C., Bove J.M.;
 RT "Organization and nucleotide sequences of the Spiroplasma citri genes
 for ribosomal protein S2, elongation factor Ts, spiralin,
 phosphofructokinase, pyruvate kinase, and an unidentified protein.";
 RT J. Bacteriol. 172:2693-2703(1990).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 27556 / R8A2;
 RA le Dantec L.;
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate -> ADP + D-fructose 1,6-bisphosphate.
 CC -1- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOFRUCTOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL; AF012877; AAB69998.1; -
 DR PIR; E35270; E35270.
 DR HSSP; P06998; 2PK.
 DR InterPro; IPR000023; pfrrckinase.
 DR Pfam; PF00365; PKF; 1.
 DR PRINTS; PR00476; PFRCKTKINASE.
 DR PRODOM; PD000707; Pfrckinase; 1.
 DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 1.
 KW Kinase; Transferase; Glycolysis.
 SQ SEQUENCE 327 AA; 35197 MW; 19D9D47E176FE2F6 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 327;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMHQRK 10
 DB 211 VAMHQRK 1

DB 206 VAMHQRK 214

RESULT 3

ID RUVB.CHLPN STANDARD; PRT; 337 AA.
 AC Q928F3; Q9JRV2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Holliday junction DNA helicase ruvb.
 DE RUVB OR CPN0390 OR CP0365.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxId=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; Pubmed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; Pubmed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Knout H., Craven B., Bowman C., Dodson R.,
 RA Eissen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MOpn and Chlamydia
 pneumoniae AR39.";
 RT Nucleic Acids Res. 28:1397-1406(2000).
 RL [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; Pubmed=10871362;
 RA Shirai M., Hiraoka H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishi K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- FUNCTION: THE RUVA-RUVB COMPLEX IN THE PRESENCE OF ATP RENATURALIZES
 CC CRUCIFORM STRUCTURE IN SUPERCOILED DNA WITH PALINDROMIC SEQUENCE,
 CC INDICATING THAT IT MAY PROMOTE STRAND EXCHANGE REACTIONS IN
 CC HOMOLOGOUS RECOMBINATION. RUVAB IS AN HELICASE THAT MEDIATES THE
 CC HOLLIDAY JUNCTION MIGRATION BY LOCALIZED DENATURATION AND
 CC REANNEALING (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS A COMPLEX WITH RUVA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RUVA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL; AE001623; AAD18533.1; -
 DR EMBL; AE002198; AAF38214.1; -
 DR EMBL; AP002546; BAA98598.1; -
 DR TIGR; CP0365; -
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_cent.
 DR InterPro; IPR004605; RUVB.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFRAMS; TIGR00635; ruvb; 1.
 DR DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
 KW Complete proteome.

FT NP_BIND 57 64 ATP (POTENTIAL).
 FT CONFLICT 168 168 T -> A (IN REF. 1).
 SQ SEQUENCE 337 AA: 37398 MW: D591B593F342EA2F CRC64;

Query Match 66.0%; Score 33; DB 1; Length 337;
 Best Local Similarity 55.6%; Pred. No. 11;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAMHDERK 10
 DB 5 VAVLHODK 13

RESULT 4

PTPE_HUMAN STANDARD; PRT; 700 AA.
 ID PTPE_HUMAN
 AC P23469;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase epsilon precursor (EC 3.1.3.48) (R-PTP-epsilon).
 GN PTPE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=91006018; PubMed=2170109;
 RA Krueger N.X., Streuli M., Saito H.;
 RT "Structural diversity and evolution of human receptor-like protein tyrosine phosphatases."
 RL EMBO J. 9:3241-3252(1990).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O - protein tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: X54134; CA38069.1; -.
 DR PIR: S12053; S12053.
 DR HSP: P18052; IYFO.
 DR Genew: HGNC:9669; PTPRE.
 DR MIM: 600926; -.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR InterPro: IPR000242; TYR_PP.
 DR Pfam: PF00102; Y_phosphatase_2.
 DR PRINTS: PR00700; PTPYPPHASE.
 DR SMART: SM00194; PTPC; 2.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE: PS50055; TYR_PHOSPHATASE_PP; 2.
 KW Glycoprotein; Transmembrane; Hydrolase; Phosphorylation; Repeat; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 700 PROTEIN-TYROSINE PHOSPHATASE EPSILON.
 FT DOMAIN 20 46 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 47 69 POTENTIAL.
 FT DOMAIN 70 700 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 154 393 PROTEIN-TYROSINE PHOSPHATASE 1.
 FT DOMAIN 394 700 PROTEIN-TYROSINE PHOSPHATASE 2.
 FT ACT_SITE 335 335 BY SIMILARITY.
 FT ACT_SITE 630 630 BY SIMILARITY.
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 700 AA: 80641 MW: D096BCADCEA65708 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 700;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMHDERK 10
 DB 352 MAMHAEOK 360

RESULT 5

PTPE_RAT STANDARD; PRT; 796 AA.
 ID PTPE_RAT
 AC Q03348;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase alpha precursor (EC 3.1.3.48) (R-PTP-alpha).
 GN PTPE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=93038682; PubMed=1417854;
 RA Moriyma T., Fujiwara Y., Imai E., Takenaka M., Kawanishi S., Itoe T., Noguchi T., Tanaka T., Ueda N.;
 RT "cDNA cloning of rat LRP, a receptor like protein tyrosine phosphatase, and evidence for its gene regulation in cultured rat mesangial cells."
 RL Biochem. Biophys. Res. Commun. 188:34-39(1992).
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O - protein tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: L01702; AAA41983.1; -.
 DR HSP: P18052; IYFO.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR InterPro: IPR000242; TYR_PP.
 DR Pfam: PF00102; Y_phosphatase_2.
 DR PRINTS: PR00700; PTPYPPHASE.
 DR SMART: SM00194; PTPC; 2.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE: PS50055; TYR_PHOSPHATASE_PP; 2.
 KW Glycoprotein; Transmembrane; Hydrolase; Phosphorylation; Repeat; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 796 PROTEIN-TYROSINE PHOSPHATASE ALPHA.
 FT DOMAIN 20 145 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 146 169 POTENTIAL.
 FT DOMAIN 170 796 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 234 494 PROTEIN-TYROSINE PHOSPHATASE 1.
 FT DOMAIN 495 796 PROTEIN-TYROSINE PHOSPHATASE 2.
 FT ACT_SITE 436 436 BY SIMILARITY.
 FT ACT_SITE 726 726 BY SIMILARITY.
 FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 796 AA; 4793796191056920 CRC64;

Query Match
Best Local Similarity 66.0%; Score 33; DB 1; Length 796;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MMHQRK 10
DB 455 MMHSRK 461

RESULT 6
PTRA_MOUSE STANDARD; PRT; 829 AA.
ID PTRA_MOUSE 18052; Q61808;
AC 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein-tyrosine phosphatase alpha precursor (EC 3.1.3.48) (R-PTP-
alpha) (LCA-related phosphatase).
GN PTPRA OR LRP OR PRPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6 X DBA/2;
RA MEDLINE=90280391; PubMed=2162042;
RA Matthews R.J., Cahill E.D., Thomas M.L.;
RT "Identification of an additional member of the protein-tyrosine-
RT phosphatase family: evidence for alternative splicing in the tyrosine
RT phosphatase domain."
RT Proc. Natl. Acad. Sci. U.S.A. 87:4444-4448(1990).
RN [2]
RP SEQUENCE OF 358-467 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RA MEDLINE=95134232; PubMed=7832766;
RA Hendriks W., Schepens J., Brugman C., Zeeuwen P., Wieringa B.;
RT "A novel receptor-type protein tyrosine phosphatase with a single
RT catalytic domain is specifically expressed in mouse brain."
RT Biochem. J. 305:499-504(1995).
RN [3]
RP SEQUENCE OF 651-756 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RA MEDLINE=93086603; PubMed=1454056;
RA Schepens J., Zeeuwen P., Wieringa B., Hendriks W.;
RT "Structural basis for inhibition of receptor protein-tyrosine
RT phosphatase-alpha by dimerization."
RT Nature 382:555-559(1996).
RN [4]
RP NATURE 382:555-559(1996).
RT "CAVITIC ACTIVITY: Protein tyrosine phosphate + H(2)O - protein
RT tyrosine + phosphate.
RT -1- SUPRACELLULAR LOCATION: Type I membrane protein.
RT -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND
RT A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
RT -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
RT -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M36033; AAA39448.1; -
CC EMBL; M36034; AAA39449.2; -
CC EMBL; 223054; CAA80589.1; -
CC EMBL; 223055; CAA80590.1; -
CC PIR; A35501; A35501.
CC PIR; B35501; B35501.
CC PDB; 1YFO; 01-APR-97.
CC MGD; MGI:97808; Pcpa.
CC InterPro; IPR000387; TYR-phosphatase.
CC InterPro; IPR000242; TYR-PP.
CC Pfam; PF00102; Y-phosphatase.
CC PRINTS; PR00700; PRTYPHTASE.
CC SMART; SM00194; PTPc; 2.
CC PROSITE; PS00383; TYR-PHOSPHATASE_1; 2.
CC PROSITE; PS00387; TYR-PHOSPHATASE_2; 2.
CC PROSITE; PS0056; TYR-PHOSPHATASE_PTP; 2.
CC Glycoprotein; Transmembrane; Hydroxylase; Phosphorylation; Signal;
CC Repeat; Alternative splicing; 3D-structure.
CC SIGNAL 1 19
CC CHAIN 20 829
CC DOMAIN 20 142
CC TRANSMEM 143 166
CC DOMAIN 232 527
CC DOMAIN 528 829
CC ACT_SITE 469 469
CC ACT_SITE 759 759
CC CARBOHYD 21 21
CC CARBOHYD 47 47
CC CARBOHYD 51 51
CC CARBOHYD 68 68
CC CARBOHYD 80 80
CC CARBOHYD 86 86
CC CARBOHYD 104 104
CC CARBOHYD 124 124
CC VARSLIC 268 303
SQ SEQUENCE 829 AA; 93697 MW; 7B1E335D4CCE09B CRC64;

Query Match
Best Local Similarity 66.0%; Score 33; DB 1; Length 829;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MMHQRK 10
DB 488 MMHSRK 494

RESULT 7
ICEL_SPOFR STANDARD; PRT; 299 AA.
ID ICEL_SPOFR 89116;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Caspase-1 precursor (EC 3.4.22.-).
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Diptera; Noctuidae; Noctuidae; Amphipyritinae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC MEDLINE=97153084; PubMed=8999805;
RA Ahmad M., Srinivasula S.M., Wang L., Litwack G., Fernandes-Alnemri T.,
RA Alnemri E.S.;
RT "Spodoptera frugiperda caspase-1, a novel insect death protease that
RT cleaves the nuclear immunophilin FKBP46, is the target of the

```

```

RT baculovirus antiapoptotic protein p35."
RL J. Biol. Chem. 272:1421-1424(1997).
CC -1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES
CC RESPONSIBLE FOR APOPTOSIS EXECUTION (BY SIMILARITY). INHIBITED BY
CC THE BACULOVIRUS ANTI-APOPTOTIC PROTEIN P35. CLEAVES P35 AND
CC NUCLEAR IMMUNOPHILIN FKBP46.
CC -1- SUBUNIT: HETERODIMER OF A 19/18 kDa (P19/P18) AND A 12 kDa (P12)
CC SUBUNIT.
CC -1- PTM: AN AUTOCATALYTIC MECHANISM GENERATES THE TWO ACTIVE SUBUNITS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U81510; AAC47442.1; -.
DR HSSP: PA2574; ICP3.
DR MEROPS: C14.015; -.
DR InterPro: IPR002398; ICE.
DR InterPro: IPR002138; ICE_P10.
DR InterPro: IPR001309; ICE_P20.
DR Pfam: PF00655; ICE_P10; 1.
DR Pfam: PF00656; ICE_P20; 1.
DR SMART: SM00115; CASC; 1.
DR PROSITE: PS01122; CASPASE_CYS; 1.
DR PROSITE: PS01121; CASPASE_HIS; 1.
DR PROSITE: PS50207; CASPASE_P10; 1.
DR PROSITE: PS50208; CASPASE_P20; 1.
KW Hydrolase; Thiol protease; zymogen; Apoptosis.
FT PROPEP 1 28
FT CHAIN 29 184 CASPASE-1 SUBUNIT P19/18.
FT PROPEP 185 195 POTENTIAL.
FT CHAIN 196 299 CASPASE-1 SUBUNIT P12.
FT ACT_SITE 136 136 BY SIMILARITY.
FT ACT_SITE 178 178 BY SIMILARITY.
SQ SEQUENCE 299 AA; 33527 MW; 99FAFED09B04EED6 CRC64;

Query Match 64.0%; Score 32; DB 1; Length 299;
Best Local Similarity 62.5%; Pred. No. 15;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 ANMHQERK 10
Db 272 ANMHQOKO 279

RESULT 8
KSF5_ECOLI STANDARD; PRT; 327 AA.
AC Q47334;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Polysialic acid capsule expression protein kpsF.
GN kpsF.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K5;
RX MEDLINE=97086613; PubMed=8932302;
RA Simpson D.A., Hammarion T.C., Roberts I.S.;
RT "transcriptional organization and regulation of expression of region 1
RT of the Escherichia coli K5 capsule gene cluster."
RL J. Bacteriol. 178:6466-6474(1996).
CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. GUTQ/KPSF SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X95264; CA64561.1; -.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR004800; kpsF.
DR InterPro: IPR001347; SIS.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF01380; SIS; 1.
DR SMART: SM00116; CBS; 2.
DR TIGRfams: TIGR00393; kpsF; 1.
DR ATP-Binding; Repeat; CBS domain.
KW NP_BIND 63 68 ATP (POTENTIAL).
FT DOMAIN 215 268 CBS 1.
FT DOMAIN 280 327 CBS 2.
SQ SEQUENCE 327 AA; 35563 MW; D0C9A38DFB1E02C4 CRC64;

Query Match 64.0%; Score 32; DB 1; Length 327;
Best Local Similarity 55.6%; Pred. No. 17;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ANMHQER 9
Db 182 AIAIHQK 190

RESULT 9
CHOR_BACTN STANDARD; PRT; 401 AA.
ID CHOR_BACTN 002550;
AC Q02550;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Chondro-6-sulfatase regulatory protein.
DE CHOR.
GN Bacteroides thetaiotaomicron.
OS Bacteroides; Bacteroidetes; Bacteroides; Bacteroidaceae;
OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidaceae;
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=46-1, and CS3;
RX MEDLINE=93054330; PubMed=1429442;
RA Cheng Q., Hwa V., Salyers A.A.;
RT "A locus that contributes to colonization of the intestinal tract by
RT Bacteroides thetaiotaomicron contains a single regulatory gene (chor)
RT that links two polysaccharide utilization pathways."
RL J. Bacteriol. 174:7185-7193(1992).
CC -1- FUNCTION: LINKS THE HEPARIN AND THE CHONDROITIN SULFATE
CC UTILIZATION PATHWAYS WHICH CONTRIBUTE TO THE COLONIZATION OF THE
CC INTESTINAL TRACT. MAY HAVE A REGULATOR FUNCTION. REQUIRED FOR
CC THE EXPRESSION OF CHONDRO-6-SULFATASE.
CC -1- SIMILARITY: BELONGS TO THE ASLB/ATSB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L00678; AAA22908.1; -.
DR PIR: A47014; A47014.
KW Transcription regulation; Activator.
SQ SEQUENCE 401 AA; 46432 MW; 43C4BE4018AA3735 CRC64;

```

Query Match 64.0%; Score 32; DB 1; Length 401;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

2 VAMMOERK 10
 316 VEMMSERO 324

RESULT 10
 ANA_DROME STANDARD; PRT; 474 AA.

AC 026307; 09Y524; 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Anachronism protein precursor.
 GN ANA OR CG8084.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Eye imaginal disk;
 RX MEDLINE=93327423; PubMed=791657;
 RA Ebens A.J., Garren H., Chevette B.N.R., Zipursky S.L.;
 RT "The Drosophila anachronism locus: a glycoprotein secreted by glia
 RT inhibits neuroblast proliferation.";
 RL Cell 74:15-27(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazek R.G.,
 RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummel S.R., Kaira K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomtan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Pouch E.,
 RA Sequiera A., Sethi H., Sitr E., Svirska R.R., Twomey B., Wan K.H.,
 RA Weiburg T., Zhang R., Zheran L.L., Rubin G.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBD databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: NEGATIVELY REGULATES PROLIFERATION OF NEURONAL PRECURSOR
 CC CELLS, THEREBY CONTROLLING THE TIMING OF POSTEMBRYONIC
 CC NEUROGENESIS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN SOME GLIAL CELLS AND SECRETED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S63815; AAB27582.1; -
 DR EMBL; AC006074; -; NOT ANNOTATED_CDS.
 DR EMBL; AF003834; AF58998.1; -
 DR Flybase; FBgn0011746; ana.
 KW Glycoprotein; Signal; Developmental protein; Neurogenesis.
 FT SIGNAL 1 33
 FT CHAIN 34 474
 FT DOMAIN 453 462
 FT CARBOHYD 54 54
 FT CARBOHYD 62 62
 FT CARBOHYD 73 73
 FT CARBOHYD 116 116
 FT CARBOHYD 144 144
 FT CARBOHYD 342 342
 FT CONFLICT 368 388
 FT SEQUENCE 474 AA; 54015 MW; FE8A43042B595A1 CR664;
 SQ

Query Match 64.0%; Score 32; DB 1; Length 474;
 Best Local Similarity 55.6%; Pred. No. 26;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

2 VAMMOERK 10
 216 VSVIHQOR 224

RESULT 11
 MG1A_TREPA STANDARD; PRT; 496 AA.

AC 056342;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Galactoside transport ATP-binding protein mg1a homolog.
 GN MG1A OR TP0685.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_Taxid=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97080510; PubMed=8921855;
 RA Porcella S.F., Popova T.G., Hagman K.E., Penn C.W., Radolf J.D.,
 RA Nordard M.V.;
 RT "A mg1-like operon in Treponema pallidum, the syphilis spirochete.";
 RL Gene 177:115-121(1996).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weldman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 RT spirochete.";
 RL Science 281:375-388(1998).
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 CC FOR GALACTOSIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO
 CC THE TRANSPORT SYSTEM (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (Potential).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MGLA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U48416; AAC44585.1; -
 CC EMBL: AE001242; AAC65648.1; -
 CC TRIGR: TP00685; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR Transport: Sugar transport; Membrane; ATP-binding; Complete proteome.
 KW NE_BIND 37 44 ATP (POTENTIAL).
 FT SEQUENCE 496 AA; 55191 MW; B6F63D53C5CD111 CRC64;
 SQ

Query Match 64.0%; Score 32; DB 1; Length 496;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VAMH0E 8
 |||:|
 Db 81 VAMV0E 87

RESULT 12
 FRDA_WOLSU STANDARD; PRT; 656 AA.
 ID FRDA_WOLSU
 AC P17412;
 DT 01-NOV-1990 (Rel. 15, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fumarate reductase flavoprotein subunit (EC 1.3.99.1).
 GN FRDA.
 OS Wolinella succinogenes.
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Wolinella.
 OX NCBI_TaxID=844;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91058386; PubMed=2244791;
 RA Lauterbach F., Koertner C., Albracht S.P., Unden G., Kroeger A.;
 RT "The fumarate reductase operon of Wolinella succinogenes. Sequence
 RT and expression of the frda and frdb genes.";
 RL Arch. Microbiol. 154:386-393(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Simon J., Gross R., Ringel M., Schmidt E., Kroeger A.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), AND REVISIONS TO 281-289.
 RX MEDLINE=20052158; PubMed=10586875;
 RA Lancaster C.R.D., Kroeger A., Auer M., Michel H.;
 RT "Structure of fumarate reductase from Wolinella succinogenes at 2.2 A
 RT resolution.";
 RL Nature 402:377-385(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).
 RX MEDLINE=21145485; PubMed=11248702;
 RA Lancaster C.R.D., Gross R., Simon J.;
 RT "A third crystal form of Wolinella succinogenes quinol:fumarate
 RT reductase reveals domain closure at the site of fumarate reduction.";
 RL Eur. J. Biochem. 268:1820-1827(2001).
 CC -1- CATALYTIC ACTIVITY: Succinate + acceptor = fumarate + reduced
 CC acceptor.
 CC -1- COFACTOR: FAD; COVALENTLY LINKED.
 CC -1- SUBUNIT: FUMARATE DEHYDROGENASE FORMS PART OF AN ENZYME COMPLEX
 CC CONTAINING THREE SUBUNITS: A FLAVOPROTEIN, AN IRON-SULFUR, AND
 CC CYTOCHROME B-556.
 CC -1- SIMILARITY: TO THE FLAVOPROTEIN SUBUNITS OF OTHER SPECIES
 CC FUMARATE REDUCTASE AND OF SUCCINATE DEHYDROGENASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ000662; CA04214.1; -
 CC PIR: S10165; S10165.
 DR PIR: B44954; B44954.
 DR PDB: 1QEA; 29-NOV-99.
 DR PDB: 1QEB; 29-NOV-99.
 DR InterPro: IPR003953; FAD_bind2.
 DR InterPro: IPR003952; FRD/SDH_FAD.
 DR InterPro: IPR004112; Succ_DH_flav_C.
 DR Pfam: PF00890; FAD-binding_2; 1.
 DR PROSITE: PS00504; FRD_SDH_FAD_BINDING; 1.
 DR OXidoreductase; Electron transport; Flavoprotein; FAD; 3D-structure.
 KW NE_BIND 9 23 FAD (COVALENT).
 FT BINDING 43 43 FAD (COVALENT).
 FT ACT_SITE 257 257 FAD (COVALENT).
 FT ACT_SITE 273 273 BY SIMILARITY.
 FT CONFLICT 281 289 BY SIMILARITY.
 FT SEQUENCE 656 AA; 72834 MW; DF4D499571412326 CRC64;
 SQ

Query Match 64.0%; Score 32; DB 1; Length 656;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AVAMH0ERK 10
 |:|:|:|
 Db 181 AIALIHQDK 190

RESULT 13
 PTR_A_HUMAN STANDARD; PRT; 802 AA.
 ID PTR_A_HUMAN
 AC P18433; Q14513;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase alpha precursor (EC 3.1.3.48) (R-PTP-
 DE alpha).
 GN PTPRA OR PTPA.
 GN Homo sapiens (human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-90349565; PubMed2166945;
 RA Sap J., D'Eustachio P., Givol D., Schlessinger J.;
 RT "Cloning and expression of a widely expressed receptor tyrosine
 RT phosphatase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6112-6116(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90384936; PubMed2169617;
 RA Kaplan R., Morse B., Huebner K., Croce C., Hawk R., Ravera M.,
 RA Rice G., Jaye M., Schlessinger J.;
 RT "Cloning of three human tyrosine phosphatases reveals a multigene
 RT family of receptor-linked protein-tyrosine-phosphatases expressed in
 RT brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:7000-7004(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91006018; PubMed2170109;
 RA Krueger N.X., Streuli M., Saito H.;
 RT "Structural diversity and evolution of human receptor-like protein
 RT tyrosine phosphatases.";
 RL EMBO J. 9:3241-3252(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Kidney;
 RA MEDLINE-91088320; PubMed2175890;
 RA Ohagi S., Nishi M., Steiner D.F.;
 RT "Sequence of a cDNA encoding human LRP (leukocyte common antigen-
 RT related peptide).";
 RL Nucleic Acids Res. 18:7159-7159(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91032191; PubMed2172030;
 RA Jirik F.R., Janzen N.M., Melhado I.G., Harder K.W.;
 RT "Cloning and chromosomal assignment of a widely expressed human
 RT receptor-like protein-tyrosine phosphatase.";
 RL FEBS Lett. 273:239-242(1990).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21638749; PubMed11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharajallo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillips B.J.C.T., Pirithalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [7]
 RP CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O -> protein
 CC tyrosine + phosphate.
 CC [1]. SUBCELLULAR LOCATION: Type I membrane protein.
 CC [1]. SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34668; AAA36528.1; -
 DR EMBL; X54130; CAA38065.1; -
 DR EMBL; X54890; CAA38662.1; -
 DR EMBL; X53364; CAA37447.1; -
 DR EMBL; AL121905; CAC10337.1; -
 DR PIR; A36065; A36065.
 DR PIR; S12049; S12049.
 DR HSSP; P18052; LYPO.
 DR GeneW; HGNC:9664; PTPRA.
 DR MIM; 176884; -
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; TYR_PP.
 DR Pfam; PF00102; Y_phosphatase; 2.
 DR PRINTS; PR00700; PRTYPPHTASE.
 DR SMART; SM00194; PTPC; 2.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 2.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 2.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 2.
 KW Glycoprotein; Transmembrane; Hydrolase; Phosphorylation; Signal;
 KW Repeat; Alternative splicing.
 FT FT SIGNAL 1 19
 FT CHAIN 802
 FT DOMAIN 20 142
 FT TRANSMEM 143 165
 FT DOMAIN 166 802
 FT DOMAIN 241 500
 FT DOMAIN 501 802
 FT ACT_SITE 442 442
 FT ACT_SITE 732 732
 FT CARBOHYD 21 21
 FT CARBOHYD 21 21
 FT CARBOHYD 36 36
 FT CARBOHYD 68 68
 FT CARBOHYD 80 80
 FT CARBOHYD 86 86
 FT CARBOHYD 104 104
 FT CARBOHYD 124 124
 FT CARBOHYD 139 147
 FT VARSPLIC 124 147
 FT CONFLICT 114 114
 FT CONFLICT 112 112
 FT CONFLICT 122 122
 FT CONFLICT 138 138
 FT CONFLICT 179 187
 FT CONFLICT 269 289
 FT CONFLICT 367 367
 FT CONFLICT 493 493
 FT CONFLICT 786 786
 SQ SEQUENCE 802 AA; 90599 MW; 8B964C3B5B5B32 CRC64;
 Query Match 64.0%; Score 32; DB 1; Length 802;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 4 MMH0ERK 10
 DB 461 MMH0ERK 467
 RESULT 14
 ENTK_BOVIN
 AC P98072; STANDARD; PRT; 1035 AA.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enteroprotease precursor (EC 3.4.21.9) (Enterokinase).
 GN PRS7 OR ENTK.
 OS Bos taurus (Bovine).

Query Match 64.0%; Score 32; DB 1; Length 1035;
 Best Local Similarity 66.7%; Pred. No. 61;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VAMMHOERK 10
 DB 893 IAMHLEMK 901

RESULT 15

HEM_DROME
 ID HEM_DROME STANDARD; PRT; 1126 AA.
 AC P53162; Q9YNU8;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Membrane-associated protein Hem (Dhem-2).
 GN HEM OR HEM2 OR CG5837.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Canton-S;
 RX MEDLINE=9531111; PubMed=7643388;
 RA Baumgartner S., Martin D., Chiquet-Ehrismann R., Sutton J.,
 RA Desai A., Huang I., Kato K., Hromas R.,
 RT "The HEM proteins: a novel family of tissue-specific transmembrane
 RT proteins expressed from invertebrates through mammals with an
 RT essential function in oogenesis.";
 RL J. Mol. Biol. 251:41-49(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harits N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon S., Nusskern D.R., Pacleib J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: PLAY A ROLE DURING GROWTH OF THE OOCYTE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED MATERNALLY IN THE OOCYTE AND SHOWS
 CC UNIFORM EXPRESSION DURING THE FIRST HALF OF EMBRYOGENESIS, BUT BECOMES
 CC BECOMES RESTRICTED TO THE BRAIN AND THE NERVOUS SYSTEM DURING LATE
 CC EMBRYOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE HEM-1/HEM-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X80028; CAA56332.1; -;
 DR EMBL; AE003597; AAF51820.1; -;
 DR FlyBase; FBgn0011771; Hem.
 KW Transmembrane.
 KM TRANSMEM 989 1006
 FT SEQUENCE 1126 AA; 129379 MW; 125FE7177AECC0B5 CRC64;
 POTENTIAL.
 QY 3 AMMHOERK 10
 DB 331 ALMHRER 338

Search completed: July 11, 2003, 11:55:21
 Job time: 1.95634 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 4.3659 Seconds

(without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-9

Perfect score: 50

Sequence: 1 AVAMMHQERK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp rodent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.virus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	82.0	380	11	Q920J5
2	38	76.0	245	5	Q9NB28
3	37	74.0	380	11	Q9D695
4	36	72.0	553	2	Q93AG2
5	35	70.0	212	13	Q42614
6	35	70.0	376	5	P91376
7	35	70.0	500	13	Q8UV19
8	35	70.0	578	10	Q23030
9	35	70.0	602	10	Q22783
10	35	70.0	1561	11	Q88286
11	35	70.0	1568	4	Q95785
12	34	68.0	82	4	Q91509
13	34	68.0	82	11	Q9CR20
14	34	68.0	695	10	Q8W2R3
15	33	66.0	200	10	Q9FUS9
16	33	66.0	275	16	O50755

17	33	66.0	281	16	Q8YVK5	Q8YVK5 anabaena sp
18	33	66.0	303	2	Q9RC27	Q9RC27 bacillus sp
19	33	66.0	318	10	Q94JB7	Q94JB7 oryza sativ
20	33	66.0	349	5	Q9USN7	Q9USN7 leishmania
21	33	66.0	366	16	Q98523	Q98523 rhizobium 1
22	33	66.0	395	13	Q90231	Q90231 alligator m
23	33	66.0	536	4	Q8TE48	Q8TE48 homo sapien
24	33	66.0	642	4	Q96K06	Q96K06 homo sapien
25	33	66.0	761	5	Q95Y05	Q95Y05 leishmania
26	33	66.0	793	11	Q91V35	Q91V35 mus muscula
27	33	66.0	1018	5	Q95TW8	Q95TW8 drosophila
28	33	66.0	1232	5	Q9VFI9	Q9VFI9 drosophila
29	32	64.0	161	10	Q8VWR4	Q8VWR4 arabidopsis
30	32	64.0	229	10	Q9FZJ8	Q9FZJ8 arabidopsis
31	32	64.0	252	16	Q8YGAS	Q8YGAS brucella me
32	32	64.0	298	16	P71702	P71702 mycobacteri
33	32	64.0	473	13	Q90ZL9	Q90ZL9 xenopus lae
34	32	64.0	473	13	Q91B97	Q91B97 potamogetrygo
35	32	64.0	486	16	Q91278	Q91278 pseudomonas
36	32	64.0	496	2	Q56341	Q56341 treponema p
37	32	64.0	497	2	Q9ZGA2	Q9ZGA2 treponema d
38	32	64.0	526	5	Q8T823	Q8T823 drosophila
39	32	64.0	561	5	Q9VNM4	Q9VNM4 drosophila
40	32	64.0	708	4	Q96S08	Q96S08 homo sapien
41	32	64.0	738	10	Q9XEM3	Q9XEM3 oryza sativ
42	32	64.0	802	4	Q96TDP	Q96TDP homo sapien
43	32	64.0	807	13	Q91969	Q91969 gallus gall
44	32	64.0	1220	4	Q9P252	Q9P252 homo sapien
45	32	64.0	1972	4	Q9UIF8	Q9UIF8 homo sapien

ALIGNMENTS

RESULT 1	ID	Q920J5	PRELIMINARY:	PRT:	380 AA.
AC	Q920J5	01-DEC-2001 (TREMURel. 19, Created)			
DT	01-DEC-2001 (TREMURel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMURel. 20, Last annotation update)				
DE	Meglin.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-2136806; PubMed-11473647;				
RA	Nagaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,				
RA	Iagi M., Nagano N., Inagi R., Kurokawa K.;				
RT	"Cloning of rodent meglin revealed its up-regulation in				
RT	mesangiolipoliferative nephritis."				
RL	Kidney Int. 60:641-652(2001).				
CC	-1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.				
DR	EMBL: AF105329; AAL16769.1; -				
DR	InterPro: IPR000215; Serpin.				
DR	Pfam: PF00079; serpin.1.				
DR	PROSITE: PS00284; SERPIN; UNKNOWN_1.				
KW	Serpin.				
SO	SEQUENCE 380 AA; 42821 MW; D8076CABEE2C2FC CRC64;				
QY	Query Match	82.0%;	Score 41;	DB 11;	Length 380;
DB	Best Local Similarity	80.0%;	Pred. No. 1.9;		
	Matches 8; Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;	
	1 AVAMMHQERK 10				
	204 AVAMMHQERR 213				
RESULT 2					
Q9NB28					

```

ID 09NB28 PRELIMINARY; PRT; 245 AA.
AC 09NB28;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HMG box transcription factor Amphiox/2/3.
OC Branchiostoma floridae (Florida lancelet) (Amphioxus);
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RA Holland L.Z., Schubert M., Holland N.D., Neuman T.;
RT "Evolutionary conservation of the presumptive neural plate markers
RT Amphiox/2/3 and AmphibNeurogenin in the invertebrate chordate
RT amphioxus."
RL Dev. Biol. 0:0-0(2000).
DR EMBL: AF271787; AAF81765.1; -.
DR HSSP: Q05066; 1HRX.
DR InterPro: IPR000135; Highmobily_12.
DR InterPro: IPR000910; HMG_12_box.
DR Pfam: PF00505; HMG_box.1.
DR PRINTS: PR00086; HIGHMOBLTY12.
DR SMART: SM00398; HMG; 1.
SQ SEQUENCE 245 AA; 28158 MW; F4B58A2P4C162299 CRC64;

```

```

Query Match 76.0%; Score 38; DB 5; Length 245;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 4 MMHGERK 10
DB 6 MMHGERK 12

RESULT 3
OY 09D695 PRELIMINARY; PRT; 380 AA.
ID 09D695;
AC 09D695;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 4631416M05Rik protein (Megasin).
GN SERPINB7 OR 4631416M05Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SKIN.
RX MEDLINE=21085660; PubMed=11217851;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasakii H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; PubMed=11473647;
RA Nagaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangio proliferative nephritis."
RL Kidney Int. 60:641-652(2001).
DR EMBL: AK014524; BAB29410.1; -.
DR EMBL: AF105328; AAL16768.1; -.
DR HSSP: P05121; 1ATC.
DR MGD: MGI:2151053; Serpinb7.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CFA CRC64;

```

```

Query Match 74.0%; Score 37; DB 11; Length 380;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 VAMMGERK 10
DB 205 VAMMGERK 213

RESULT 4
OY 093AG2 PRELIMINARY; PRT; 553 AA.
ID 093AG2;
AC 093AG2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE FAD-monoxygenase.
GN CAD.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FO07;
RA trevino-Quintanilla L.G., Soberon-Chavez G.;
RT "Isolation and characterization of a Pseudomonas putida strain able to
RT degrade chloranilic acid (2,5-dichloro-3,6-dihydroxy benzo-1,4-
RT quinone)."
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF421356; AAL16082.1; -.
DR InterPro: IPR000733; Flay_monoxygenase.
DR InterPro: IPR002938; MoxY_FAD_binding.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; Thle_domain.
DR Pfam: PF01494; FAD_binding-3; 1.
DR Pfam: PF01360; Monoxygenase; 1.
KW Monoxygenase.
SQ SEQUENCE 553 AA; 61430 MW; 54BB707D639E0C6E CRC64;

Query Match 72.0%; Score 36; DB 2; Length 553;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

DE Antithrombin III (Fragment).
 OS Cyprinus carpio (Common carp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 RX NCBI_Taxid=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gracey A.Y., Colonna-Romano S., Tjku P.E., Cossins A.R., Maresca B.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL; AF006495; AAC61755.1; -;
 DR EMBL; AF006494; AAC19409.1; -;
 DR HSSP; P01008; 1BR8.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
 KW Serpin.
 FT NON_TER
 SQ SEQUENCE 212 AA; 23875 MW; 7F185103CF8BAFF0 CRC64;
 Query Match 70.0%; Score 35; DB 13; Length 212;
 Best Local Similarity 77.8%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 VAMMHQERK 10
 DB 27 VPMVQERK 35
 ID P91376 PRELIMINARY; PRT; 376 AA.
 AC P91376;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 42.4 kDa protein.
 GN K11H12.9
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 RX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Bradshaw H.;
 RT "The sequence of C. elegans cosmid K11H12.";
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88168; AAC24403.2; -;
 DR InterPro: IPR000719; Euk_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR PRODOM: PD000001; Euk_pkinase.2.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 KW Hypothetical protein.
 SQ SEQUENCE 376 AA; 42391 MW; 60F460A4685000F8 CRC64;

Query Match 70.0%; Score 35; DB 5; Length 376;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AVAMMHQ 7
 DB 208 AVAMMHQ 214
 ID O8UV19 PRELIMINARY; PRT; 500 AA.
 AC O8UV19;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 63.6 kDa protein (Fragment).
 GN C16ORF35.
 OS Sphaeroidea nehelus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Sphaeroidea.
 RX NCBI_Taxid=39110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21096913; PubMed=11157800;
 RA Flint J., Tufarelli C., Peden J., Clark K., Daniels R.J., Hardison R.,
 RA Miller W., Phillips S., Tan-Un K.C., McMorro T., Frampton J.,
 RA Alter B.P., Frischauf A.M., Higgs D.R.;
 RT "Comparative genome analysis delimits a chromosomal domain and
 RT identifies key regulatory elements in the alpha globin cluster.";
 RL Hum. Mol. Genet. 10:371-382(2001).
 DR EMBL; AY016023; AAL73050.1; -;
 DR InterPro: IPR005365; UPF0171.
 DR Pfam: PF03666; UPF0171.1.
 FT NON_TER
 SQ SEQUENCE 500 AA; 56492 MW; 8545420A768B1879 CRC64;
 Query Match 70.0%; Score 35; DB 13; Length 500;
 Best Local Similarity 50.0%; Pred. No. 45;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AVAMMHQERK 10
 DB 89 AVALQHEERR 98
 ID O23030 PRELIMINARY; PRT; 578 AA.
 AC O23030;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TIG11.21 protein.
 GN TIG11.21.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eufrosidites II; Brassicales; Brassicaceae; Arabidopsis.
 RX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Osborne B.I., Vysotskaya V.S., Toriumi M., Yu G., Ofl O., Liu S.,
 RA Li J., Hoang L., Buehler E., Conway A.B., Conway A.R., Dewar K.,
 RA Feng J., Kim C., Kurtz D., Li Y., Shinn P., Sun H., Davis R.W.,
 RA Ecker J.R., Federspiel N.A., Theologis A.;
 RT "The sequence of BAC TIG11 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.


```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Theologis A.;
RL Submitted (JOC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: ACO02376; AAB80639.1; -.
DR InterPro: IPR000008; C2.
DR Pfam: PF00168; C2; 1.
DR SMART: SM00239; C2; 1.
SQ SEQUENCE 578 AA; 64860 MW; 5C3D8369A580C9D2 CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 10; Length 578;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVAMHQRK 10
DB 29 AVAMVHSRK 38
|||||

RESULT 9
ID 022783 PRELIMINARY; PRT; 602 AA.
AC 022783;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Atg933320 protein.
GN Atg933320.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA "MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Freus D., Niemman W.C., White O., Eisen J.A.,
RA Salzman S.L., Fraser C.M., Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RA thaliana."
RT Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: ACO02332; AAB80650.1; -.
DR InterPro: IPR000008; C2.
DR Pfam: PF00168; C2; 1.
DR SMART: SM00239; C2; 1.
SQ SEQUENCE 602 AA; 66580 MW; 8045E5A2B895C26A CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 10; Length 602;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVAMHQRK 10
DB 31 AVAMVHSRK 40
|||||

RESULT 10
ID 088286 PRELIMINARY; PRT; 1561 AA.
AC 088286;
DT 01-NOV-1998 (TREMBLrel. 08, Created)

```

```

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE WIZL.
OS WIZ OR WIZ.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Matsumoto K., Ishii N., Yoshida S., Shiosaka S., Wanaka A.,
RA Tohyama M.;
RT "Molecular Cloning and Distinct Developmental Expression Pattern of
RT Spliced Forms of A Novel Zinc Finger Gene wiz in the Cerebellum."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AB012265; BAA32790.1; -.
DR MGD; MGI:1332638; Wiz.
DR InterPro: IPR000822; znf_C2H2.
DR Pfam: PF00096; znf_C2H2; 10.
DR PRINTS: PRO0048; ZINC_FINGER.
DR SMART: SM00355; znf_C2H2; 10.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 7.
KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ SEQUENCE 1561 AA; 171243 MW; 8A010A7B5374CF6 CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 11; Length 1561;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVAMHQRK 10
DB 604 AMALKHEERK 613
|||||

RESULT 11
ID 095785 PRELIMINARY; PRT; 1568 AA.
AC 095785;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Human homolog of MUS MUSCULUS WIZL protein (Human homolog of MUS
DE MUSCULUS WIZL protein) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Gaines J.,
RA Danganan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Atlix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RT "Sequence analysis of a 3.5 Mb region in 19p13.1 between OLFR and
RT JUND."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 783-1568 FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Brower A., Ramirez M.,
RA Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A.,
RA Gaines J., Danganan L., Erlar A., Christensen M., Georgescu A.,
RA Avila J., Liu S., Atlix C., Andreise T., Trankheim M.,
RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,
RA Thomas P., Quan G., Krommiller B., Arellano A., Sanders C., Ow D.,
RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;

```


"Sequence analysis of an ~1.5 Mb contig in 19p13.1 between OLFR and
RT D19S8B5.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007059; AAC19818.1; -
DR EMBL: AC006128; AAC97985.1; -
DR EMBL: AC007059; AAC19817.1; -
DR InterPro: IPR00822; 2nf.C2H2.
DR Pfam: PF00096; zf.C2H2; 10.
DR SMART: SM00355; 2nf.C2H2; 10.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 7.
KW DNA-binding; Metal-binding; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1568 AA; 170077 MW; 7AF0D34D45F565D0 CRC64;

Query Match 70.0%; Score 35; DB 4; Length 1568;
Best Local Similarity 60.0%; Pred. No. 1,4e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AVAMHOER 10
Db 610 AMALKEER 619

RESULT 12
ID O9Y5U9 PRELIMINARY; PRT: 82 AA.
AC O9Y5U9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE HSPC039 protein (Hypothetical 9.0 kDa protein) (Immediate early
response 3 interacting protein) (Similar to HSPC039 protein).
CN DKFZPS64B1471.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye M., Zhang Q., Zhou J., Shen Y., Guan Z., Wu X., Fan H., Mao H.,
RA Dai M., Huang Q., Chen S., Chen Z.;
RT "Human HSPC039 mRNA, complete cds."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-ADRENAL GLAND;
RA Huang C., Zhang C., Wu T., Peng Y., Gu Y., Gu W., Jiang C., Li Y.,
RA Han Z., Wang Y., Chen Z., Fu G.;
RT "A novel gene expressed in human adrenal gland."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Bloecher H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RA Yiu W.H., Kok L.D.S., Fung K.P., Lee C.Y., Tsui S.K.W., Waye M.M.Y.;
RT "Characterization of immediate early response 3 interacting protein
(IERIP)."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE-MYELOID LEUKEMIA CELLS;
RA Strausberg R.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF125100; AAD39917.1; -
DR EMBL: AF164798; AAF80762.1; -
DR EMBL: AL136667; CAB66602.1; -
DR EMBL: AF371963; AAK53816.1; -
DR EMBL: BC010888; AAH10888.1; -
DR EMBL: BC017391; AAH17391.1; -
KW Transmembrane.
SQ SEQUENCE 82 AA; 8969 MW; 2015B211F0AF580 CRC64;

Query Match 68.0%; Score 34; DB 4; Length 82;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AVAMHOER 9
Db 18 AIALVHEER 26

RESULT 13
ID O9CR20 PRELIMINARY; PRT: 82 AA.
AC O9CR20;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 1110057H19R1k protein (RIKEN cDNA 1110057H19 gene).
CN 1110057H19R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE, EMBRYO, AND PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Katsukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fiedischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nodone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Winking L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kontuski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK009116; BAB26081.1; -
DR EMBL: AK004281; BAB23248.1; -
DR EMBL: AK007388; BAB25005.1; -
DR EMBL: AK007862; BAB25314.1; -
DR EMBL: BC021512; AAH21512.1; -
DR MGD; MGI:1913441; 1110057H19R1K.
SQ SEQUENCE 82 AA; 9017 MW; 3015F654F1EBB080 CRC64;

Query Match 68.0%; Score 34; DB 11; Length 82;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AVAMHOER 9

Db 18 AIAVHHEER 26

RESULT 14

Q8W2R3 PRELIMINARY; PRT; 695 AA.
 AC Q8W2R3;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative wall-associated protein kinase.
 GN OSJNB0028C16.9.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RT Saski C., Henry D., Oates R., Simmons J.;
 "Rice Genomic Sequence."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC098565; AAL69427.1; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS01187; EGF_CA; UNKNOWN_1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 KW kinase.
 SQ SEQUENCE 695 AA; 77048 MW; D9DFE3A74A655A5 CRC64;
 Query Match 68.0%; Score 34; DB 10; Length 695;
 Best Local Similarity 85.7%; Pred. No. 1e+02; 0; Mismatches 0; Gaps 0;
 Matches 6; Conservative 1; Indels 0; Gaps 0;
 QY 4 MMHQERK 10
 Db 382 MMHQRRK 388

RESULT 15

Q9FU59 PRELIMINARY; PRT; 200 AA.
 AC Q9FU59;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE P0013F10.15 protein.
 GN P0013F10.15.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Erihartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippondare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0013F10.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002523; BAB17069.1; -
 SQ SEQUENCE 200 AA; 22103 MW; 01D469471C059FD8 CRC64;

Query Match 66.0%; Score 33; DB 10; Length 200;
 Best Local Similarity 60.0%; Pred. No. 48;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AVAMHQERK 10
 Db 138 AVAIHQEQE 147

Search completed: July 11, 2003, 11:59:10
 Job time : 7.3659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 218.046 Seconds
(without alignments) 232.223 Million cell updates/sec

Title: US-10-091-442-34

Perfect score: 1959

Sequence: 1 MASLAANAAEFCEFLFREMD.....FLFYRKDDIILFSGKVSCLP 380

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1959	100.0	380	AA88379	Human megakaryocyte
2	1959	100.0	380	AA108254	Human megaslin prote
3	1959	100.0	380	AA824142	Human megaslin prote
4	1959	100.0	380	AA83075	Human megaslin prote
5	1488	76.0	380	AA64286	Rat megaslin protein
6	1477	75.4	380	AA108255	Rat megaslin protein
7	1477	75.4	380	AA824150	Rat megaslin protein
8	1477	75.4	380	AA83076	Rat megaslin protein
9	1458	74.4	368	AA108256	Mouse megaslin prote
10	1458	74.4	368	AA824151	Mouse megaslin prote

11	1458	74.4	368	AA883077	Murine megaslin prot
12	796.5	40.7	379	AA15155	Human prostate ser
13	771.5	39.4	395	AA47207	Human NOVI protein
14	769	39.3	394	AB53280	Human polypeptide
15	752	38.4	390	AA825276	SCC antigen. Synt
16	752	38.4	390	AA132077	Hepatitis B virus
17	752	38.4	390	AA152928	Human SCCA1 protei
18	752	38.4	390	AA12654	Human squamous cel
19	750	38.3	390	AA132078	Hepatitis B virus
20	748	38.2	390	AA152927	Human SCCA2 protei
21	748	38.2	390	AA172655	Human squamous cel
22	744	38.0	390	AA15242	Psoriasis-like type
23	740	37.8	390	AA15241	Human prostate ser
24	738.5	37.7	617	AA15156	Human elastase inh
25	722.5	36.9	437	AA17265	Human cancer assoc
26	722.5	36.9	437	AA17265	Human cancer assoc
27	717.5	36.6	379	AA104159	Human lung cancer-
28	715.5	36.5	391	AA141029	Human lung cancer-
29	715.5	36.5	391	AA11314	Human headpin (for
30	715.5	36.5	391	AA172651	Human lung tumor
31	715.5	36.5	391	AA172651	Human lung tumor
32	715	36.5	400	AA11030	Human lung cancer-
33	715	36.5	400	AA11315	Human lung tumor
34	715	36.5	400	AA174947	Human lung cancer-
35	714.5	36.5	391	AA152926	Human hurpin prote
36	714.5	36.5	391	AA10953	Human serine prote
37	714.5	36.5	391	AA10954	Human headpin (hea
38	702	35.8	400	AA172652	Human headpin (hea
39	690	35.2	617	AA128280	Novel human diagno
40	673	34.4	382	AA12144	Plasminogen activa
41	672.5	34.3	340	AA16928	Human novel secret
42	671	34.3	392	AA12143	Plasminogen activa
43	664.5	33.9	415	AA180473	Sequence of human
44	664.5	33.9	415	AA180473	Sequence of human
45	664.5	33.9	415	AA10921	Encodes plasminoge

ALIGNMENTS

RESULT 1	AA88379	standard; Protein: 380 AA.
ID	AA88379	
XX	AA88379;	
XX	16-AUG-1994 (first entry)	
DT	16-AUG-1994 (first entry)	
XX	Human megakaryocyte differentiation factor.	
XX	Human megakaryocyte differentiation factor: MGF; thrombopoietin;	
XX	haematopoietic stimulating factor; thrombocytopoietin; platelet;	
XX	bone marrow transplantation; cancer chemotherapy.	
XX	Homo sapiens.	
XX	EP583884-A.	
XX	23-FEB-1994.	
XX	19-JUL-1993; 93EP-0305654.	
XX	17-JUL-1992; 92JP-0212305.	
XX	04-MAR-1993; 93JP-0067339.	
XX	(SUNR) SUNTORY LTD.	
XX	(TSUJ/) TSUJIMOTO M.	
XX	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;	
XX	Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;	
XX	WPI: 1994-058782/08.	
XX	N-PSDB: AA056670.	

XX New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets
 XX
 PS Claim 7; Page 30-32; 47pp; English.
 XX
 CC Human MDF can be isolated from a culture of human epidermoid
 CC carcinoma A431 cells in protein-free medium. The MDF stimulates
 CC differentiation of megakaryocytes from myeloid cells in the presence
 CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
 CC for treatment of diseases involving a decrease in platelet number
 CC (esp. thrombocytopenia) such as occurs in bone marrow
 CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kD by
 CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
 CC
 XX
 SQ Sequence 380 AA;
 Query Match 100.0%; Score 1959; DB 15; Length 380;
 Best Local Similarity 100.0%; Pred. No. 3.6e-169;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASLAANAEEFCNLFREMDNNGNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 DB 1 MASLAANAEEFCNLFREMDNNGNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 QY 61 NTASGYGNSSNSOSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 DB 61 NTASGYGNSSNSOSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 QY 121 LYAKKERVDFTHLEDTNRNINKWVENETHGKIKNVIGGGISSAVMYLVNAVYFKGK 180
 DB 121 LYAKKERVDFTHLEDTNRNINKWVENETHGKIKNVIGGGISSAVMYLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFKSPKCSGKAVAMMHQERKFNLSVIDPMSKILLELYNGGINNYVL 240
 DB 181 WQSAFTKSETINCHFKSPKCSGKAVAMMHQERKFNLSVIDPMSKILLELYNGGINNYVL 240
 QY 241 LPENDLSEIENKLTFOILMEMENTNPRMTSKYVEVEFPQFKIEKNYKQYLRALGLKDI 300
 DB 241 LPENDLSEIENKLTFOILMEMENTNPRMTSKYVEVEFPQFKIEKNYKQYLRALGLKDI 300
 QY 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVS 380
 DB 361 FLFVIRKDDIILFSGKVS 380
 RESULT 2
 AAY08254
 ID AAY08254 standard; Protein; 380 AA.
 XX
 AC AAY08254;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Human megasin protein.
 XX
 KW Megsin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 KW human; rat; murine.
 XX
 OS Homo sapiens.
 XX
 PN W09915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98MO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.

XX (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI; 1999-276983/23.
 DR N-PSDB; AAX56712.
 PS Megsin protein expressed specifically in mesangial cells
 XX
 CC Claim 1; Page 62-64; 100pp; Japanese.
 CC
 CC This invention describes the isolation of novel megasin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as Iga nephropathy.
 CC
 XX
 SQ Sequence 380 AA;
 Query Match 100.0%; Score 1959; DB 20; Length 380;
 Best Local Similarity 100.0%; Pred. No. 3.6e-169;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASLAANAEEFCNLFREMDNNGNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 DB 1 MASLAANAEEFCNLFREMDNNGNGNVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 QY 61 NTASGYGNSSNSOSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 DB 61 NTASGYGNSSNSOSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 QY 121 LYAKKERVDFTHLEDTNRNINKWVENETHGKIKNVIGGGISSAVMYLVNAVYFKGK 180
 DB 121 LYAKKERVDFTHLEDTNRNINKWVENETHGKIKNVIGGGISSAVMYLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFKSPKCSGKAVAMMHQERKFNLSVIDPMSKILLELYNGGINNYVL 240
 DB 181 WQSAFTKSETINCHFKSPKCSGKAVAMMHQERKFNLSVIDPMSKILLELYNGGINNYVL 240
 QY 241 LPENDLSEIENKLTFOILMEMENTNPRMTSKYVEVEFPQFKIEKNYKQYLRALGLKDI 300
 DB 241 LPENDLSEIENKLTFOILMEMENTNPRMTSKYVEVEFPQFKIEKNYKQYLRALGLKDI 300
 QY 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVS 380
 DB 361 FLFVIRKDDIILFSGKVS 380
 RESULT 3
 AAB24142
 ID AAB24142 standard; Protein; 380 AA.
 XX
 AC AAB24142;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Human megasin protein sequence SEQ ID NO:2.
 XX
 KW Megsin; mesangium-predominant gene; seipin regulated; nephropathy;
 KW Iga; Immunoglobulin A; detection; renal function; renal disorder;
 KW diagnosis; biological sample; blood; urine.
 XX
 OS Homo sapiens.
 XX
 PN W0200057189-A1.
 XX
 PD 28-SEP-2000.

XX 17-MAR-2000; 2000WO-JP01646.
 XX
 XX 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KURO/) KUROKAWA K.
 PA (FUSO) FUSO PHARM IND LTD.
 PA (MIYA/) MIYATA T.
 PI Miyata T:
 XX
 XX WPI: 2000-611642/58.
 DR N-PSDB: AAA99294.
 XX
 PT Evaluating renal function comprises assaying megin protein in
 PT biological sample -
 XX
 XX Example 2; Page 66-69; 93pp; Japanese.
 XX
 CC The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megin protein in biological sample. Also
 CC described are: (1) use of a anti-megin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megin protein comprising:
 CC (a) anti-megin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the human megin
 CC protein, which is given in the exemplification of the present invention.
 XX
 SO Sequence 380 AA:

Query Match 100.0%; Score 1959; DB 21; Length 380;
 Best Local Similarity 100.0%; Pred. No. 3.6e-169;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAAEFCFNLFRMDNNGNGNFFSSLSLFAALALVRLGAODDSLQIDKLLHV 60
 DB 1 MASLAANAAEFCFNLFRMDNNGNGNFFSSLSLFAALALVRLGAODDSLQIDKLLHV 60
 QY 61 NTASGYGSSNSQSGLOSLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 DB 61 NTASGYGSSNSQSGLOSLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTNHLIEDTRRNINKNVENETHGKIKNVGEGGISSAVMVLVNAVYFEGK 180
 DB 121 LYDAKVERVDFTNHLIEDTRRNINKNVENETHGKIKNVGEGGISSAVMVLVNAVYFEGK 180
 QY 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFLSVIEDPSMKILELRYNGGINMYVL 240
 DB 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFLSVIEDPSMKILELRYNGGINMYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTNPRRMTSKYVEVFFPOFKIEKNEMQYLRALGLKIDIF 300
 DB 241 LPENDLSEIENKLTFFONLMEWTNPRRMTSKYVEVFFPOFKIEKNEMQYLRALGLKIDIF 300
 QY 301 DESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIYERKOLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIYERKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSACP 380
 DB 361 FLFVIRKDDIILFSGKVSACP 380

RESULT 4
 AAB83075
 ID AAB83075 standard: Protein: 380 AA.
 XX
 AC AAB83075;
 XX

DT 10-JUL-2001 (first entry)
 XX
 XX Human megin protein.
 XX
 KW Human; megin; mesangial cell proliferative nephritis; nephrotropic;
 KW transgenic mouse; glomerular disease; animal model; drug screening.
 XX
 OS Homo sapiens.
 XX
 XX WO200124628-A1.
 XX
 XX 12-Apr-2001.
 PD
 XX 06-OCT-2000; 2000WO-JP06988.
 PE
 XX 06-OCT-1999; 99JP-0285736.
 PR
 XX (KURO/) KUROKAWA K.
 XX (MIYA/) MIYATA T.
 XX
 XX Miyata T:
 PI
 XX WPI: 2001-300136/31.
 DR
 XX N-PSDB: AAF82438.
 XX
 PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments -
 XX
 XX Example 4; Page 44-46; 62pp; Japanese.
 PS
 CC The present sequence is human megin. The human megin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 XX
 SO Sequence 380 AA:

Query Match 100.0%; Score 1959; DB 22; Length 380;
 Best Local Similarity 100.0%; Pred. No. 3.6e-169;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAAEFCFNLFRMDNNGNGNFFSSLSLFAALALVRLGAODDSLQIDKLLHV 60
 DB 1 MASLAANAAEFCFNLFRMDNNGNGNFFSSLSLFAALALVRLGAODDSLQIDKLLHV 60
 QY 61 NTASGYGSSNSQSGLOSLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 DB 61 NTASGYGSSNSQSGLOSLKRVFSDINASHKDYDLSIVNGLFAEKYVGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTNHLIEDTRRNINKNVENETHGKIKNVGEGGISSAVMVLVNAVYFEGK 180
 DB 121 LYDAKVERVDFTNHLIEDTRRNINKNVENETHGKIKNVGEGGISSAVMVLVNAVYFEGK 180
 QY 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFLSVIEDPSMKILELRYNGGINMYVL 240
 DB 181 WQSAFTKSETINCHFKSPKCGKAVAMHQRKFLSVIEDPSMKILELRYNGGINMYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTNPRRMTSKYVEVFFPOFKIEKNEMQYLRALGLKIDIF 300
 DB 241 LPENDLSEIENKLTFFONLMEWTNPRRMTSKYVEVFFPOFKIEKNEMQYLRALGLKIDIF 300
 QY 301 DESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIYERKOLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIYERKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSACP 380
 DB 361 FLFVIRKDDIILFSGKVSACP 380

xx	RESULT 5
xx	AAG64286
xx	ID AAG64286 standard; Protein; 380 AA.
xx	AC AAG64286;
xx	DT 21-SEP-2001 (first entry)
xx	DE Rat megin protein.
xx	KW Rat; megin; renal mesangial cell; mesangium proliferative nephritis.
xx	OS Rattus norvegicus.
xx	PN WO200148019-AI.
xx	PD 05-JUL-2001.
xx	PF 26-DEC-2000; 2000WO-JP09251.
xx	PR 28-DEC-1999; 9JUP-0373677.
xx	PA (KURO/) KUROKAWA K.
xx	PA (MIYA/) MIYATA T.
xx	PI Miyata T;
xx	DR WPI: 2001-425651/45.
xx	DR N-PSDB: AAH48181.
xx	PT New antibody recognizing a partial sequence of rat megin protein for
xx	PT diagnosis of mesangium proliferative nephritis -
xx	PS Disclosure: Page 54--56; 63pp; Japanese.
xx	CC The present invention relates to a novel antibody which recognises a
xx	CC peptide consisting of residues 341-354 of rat megin protein. The present
xx	CC sequence is the protein sequence for rat megin, which was used in the
xx	CC present invention. Megin is highly expressed in renal mesangial cells
xx	CC and its level is elevated in mesangium proliferative nephritis. Assay of
xx	CC the serum or urine level using the antibody is therefore indicative of
xx	CC this type of disorder.
SQ	Sequence 380 AA;
xx	
xx	Quality Match 76.0%; Score 1488; DB 22; Length 380;
xx	Best Local Similarity 73.7%; Pred. No. 2,1e-126;
xx	Matches 280; Conservative 57; Mismatches 43; Indels 0; Gaps 0

[illegible]

Db	301	VESADLGSAGSGRLYEKLMMKSLIEVSEEGTATAATSNIVYKLLPESTVFRADRP	360
QY	361	FLFVIRKDDILLFSGVKVSCP	380
		:	
Db	361	FLFVIRKNGIILFTGKRVSCP	380
RESULT 6			
AAy08255			
ID	AAy08255	standard; Protein; 380 AA.	
XX			
AC	AAy08255;		
XX			
DT	14-JUL-1999	(first entry)	
XX			
DE	Rat megalin protein.		
XX			
KW	Megalin: mesangial cell; treatment; diagnosis; disease; Iga nephropathy;		
RV	human; rat; murine.		
XX			
OS	Rattus rattus.		
XX			
PN	W09915652-A1.		
XX			
PD	01-APR-1999.		
XX			
PF	22-SEP-1998; 98WO-JP04269.		
XX			
PR	22-SEP-1997; 97JP-0275302.		
XX			
PA	(KURO/) KUROKAWA K.		
PA	(MIYA/) MIYATA T.		
XX			
PI	Miyata T;		
XX			
DR	WPI; 1999-276983/23.		
DR	N-PSDB; AAX56712.		
XX			
PT	Megalin protein expressed specifically in mesangial cells		
XX			
PS	Claim 1; Page 69-72; 100pp; Japanese.		
XX			
CC	This invention describes the isolation of novel megalin nucleic acid and		
CC	proteins from human, rat and mouse tissue. This protein is expressed		
CC	specifically in mesangial cells. The products of the invention are		
CC	useful for the treatment and diagnosis of diseases involving mesangial		
CC	cells, such as Iga nephropathy.		
XX			
Sequence	380 AA;		
50			

[illegible]

```

OY 301 DESKADLSGISGSGGLYSRMHMSXYIVTEEGTATATATGSGNIVEXKLPPOSTLRADP 360
Db 301 VESRADLSGISGGRGLYSKMLHMSLIEVSEEGTATATGSGNIVEXKLPPOSTLRADP 360
OY 361 FLFWIRKDDITILFSGKVS CP 380
Db 361 FLFWIRKNGIILFTGKVS CP 380

RESULT 7
AAB24150
ID AAB24150 standard; Protein; 380 AA.
XX
XX AAB24150:
XX
XX 30-JAN-2001 (first entry)
XX
XX Rat mesgin protein sequence SEQ ID NO:19.
XX
XX Mesgin; mesangium-predominant gene; serpin regulated; nephropathy;
KW IGA; immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
XX Rattus norvegicus.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 51
FT /note= "unspecified"
FT Misc-difference 94
FT /note= "unspecified"
FN
FN WO200057189-A1.
XX
XX 28-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-JP01646.
XX
XX 19-MAR-1999; 99JP-0075305.
XX 28-OCT-1999; 99JP-0306623.
XX
XX (KUROO/) KUROKAWA K.
PA (FUSO/) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
PA
PI Miyata T;
DR WPI; 2000-611642/58.
DR N-PSDB; AAC55238.
XX
XX Evaluating renal function comprises assaying mesgin protein in
XX biological sample -
XX
XX Example 2: Page 81-84; 93pp; Japanese.
XX
XX The present invention describes a method for evaluating renal function.
XX The method comprises assaying mesgin protein in biological sample. Also
XX described are: (1) use of a anti-mesgin protein antibody for diagnosing
XX renal function; and (2) a kit for detecting mesgin protein comprising:
XX (a) anti-mesgin protein antibody attached to solid magnetic particles;
XX (b) direct or indirect fixing for the antibody to the particles; and
XX (c) a magnet. The process is useful for evaluating renal function and
XX diagnosing renal disorders by assaying mesgin protein in biological
XX samples (preferably urine or blood). The process is reproducible and
XX gives accurate results. The present sequence represents the rat mesgin
XX protein, which is given in the exemplification of the present invention.
XX
XX Sequence 380 AA;

Query Match 75.4%; Score 1477; DB 21; Length 380;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

```

Oy	1	MASLAAAANAEECFNLFRMDNONGNANFFSSLSLPAALALVRLGAODDSLSQIDKLLHV	60
Dd	1	MASLAAAANAEECFDLFRMDSSQGNANFFSSLSLFTLALSLIRGARDCAKQIDKALHF	60
Oy	61	NTASGYGSSNSQSGLSQAKRVFSDINASHKDYDLSTVNGLFAEKVYGFHKDYIECAEK	120
Dd	61	ISPSQGNSSNSQGLQYOLKRVLADINSSHKDXLSTIANGVFAEKVDFHKSMECAEN	120
Oy	121	LYDAKVEYVDPTNLHEDTRRNINKKVENETHGKIRNVIGEGGISSAAYVLYNAVYFGK	180
Dd	121	LYNAKVEYVDPTNDIQETRFKINKWIEHETHOKIKRVLGDSLSAAYVLYNAVYFGK	180
Oy	181	MOSAFKSETLNCHEKSPKSGKAAYAAHMOEKRFPLSTYEDPSMLLETRYNGIMVYL	240
Dd	181	WKSATTKSDTLSCHEKSPSGPKAANMHQEHKRFPLSTYEDPMQILEQYHGGISMVIM	240
Oy	241	LPEMDLSEIEKRLTFQNLMEWTNPPRMTSKYVEVFPQFKIEKNEMKQYLRALGLDIF	300
Dd	241	LPEDDLSTIEEKLSTQNLMDMTNSKMKMSQYVNVFLPQFKIEKDEMHSILKSVLEDF	300
Oy	301	DEKADLSGIASGGRLYISRMMHKSYLEVTEBGETEATAGSNIVKOLPOSTLFRADHP	360
Dd	301	VESRADLSGIASGGRLYSKIMHKSLTIEVSEGTETATATESNIVEKLLPESTVERADRP	360
Oy	361	FLFYVRKDDIILFSGKVASCP 380	
Dd	361	FLFYVRKNGIILFTGKVASCP 380	
RESULT 8			
ID	AAB83076	standard; Protein: 380 AA.	
XX	AC	AAB83076;	
XX	DF	10-JUL-2001 (first entry)	
XX	DE	Rat megin protein.	
XX	KW	Rat; megin; mesangial cell proliferative nephritis; nephrotropic;	
KW	transgenic mouse; glomerular disease; animal model; drug screening.		
XX	OS	Rattus norvegicus.	
PN	WO200124628-A1.		
XX	PD	12-APR-2001.	
PF	06-OCT-2000; 2000WO-JP06988.		
PR	06-OCT-1999; 99JP-0285736.		
XX	PA	(KUROO/) KUROKAWA K.	
PA	(MIYA/) MIYATA T.		
XX	Miyata T;		
XX	WPI: 2001-300136/31.		
DR	N-PSDB; AAF82439.		
XX	PT	Mouse model for mesangial cell proliferative nephritis for development	
XX	and screening of new treatments		
PS	Disclosure; Page 48-50; 62pp; Japanese.		
CC	XX	The present sequence is rat megin. The human megin coding	
CC	sequence may be introduced into a mouse to produce an animal model of		
CC	mesangial cell proliferative nephritis. The symptoms include		
CC	enlargement of the mesangial base region, sedimentation of an immune		
CC	complex and an increase in mesangial cells. The animal model is useful		
CC	for analysing the pathology of chronic glomerular diseases and for		
CC	screening compositions for prevention and treatment of the diseases.		
CC	Highly uniform models can be made easily and in large numbers using		

CC this method.
 XX Sequence 380 AA;
 SQ
 Query Match 75.4%; Score 1477; DB 22; Length 380;
 Best Local Similarity 73.4%; Pred. No. 2e-125; Mismatches 45; Indels 0; Gaps 0;
 Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREKMDNOGNVFFSSLSIFAALALVRLGAODDSLQIDKLHV 60
 1 MASLAANAEEFGDLFREKMDNOGNVFFSSLSIFAALALVRLGAODDCAQIDKLHV 60
 DB 1 MASLAANAEEFGDLFREKMDNOGNVFFSSLSIFAALALVRLGAODDCAQIDKLHV 60
 QY 61 NTASGIVSSNSQSGLOSQKRVFSDINASHKDTLSIVNGLFAEKVYGFHKDYIECAEK 120
 61 NTASGIVSSNSQSGLOSQKRVFSDINASHKDTLSIVNGLFAEKVYGFHKDYIECAEK 120
 DB 61 ISPSROGSSNSQGLQYOLKRVLADINSSHKDKKLISANGVFAEKVDFHKSYMECAEN 120
 QY 121 LYPAAKVERVDFTNLEDTRRNINKVVENETHGKIKNYIGGGISSAVVLYNNAVYFKGK 180
 121 LYPAAKVERVDFTNLEDTRRNINKVVENETHGKIKNYIGGGISSAVVLYNNAVYFKGK 180
 DB 121 LYNAAKVERVDFTNLEDTRRNINKVVENETHGKIKNYIGGGISSAVVLYNNAVYFKGK 180
 QY 181 MQSAFTKSETINCHFKSPKCSGNAVAMHOERKFNLSVIDPSMKILELRYNGGINMYVL 240
 181 MQSAFTKSETINCHFKSPKCSGNAVAMHOERKFNLSVIDPSMKILELRYNGGINMYVL 240
 DB 181 WKSAFTKSDTSLCHFRSPGPKAVNMHOERFNLSVIDPSMKILELRYNGGINMYVL 240
 QY 241 LPENDUSEINENKLTFOMLMNTNPRRTSKYVEYFPOFKIEKNYEMKQYLALGLKDI 300
 241 LPENDUSEINENKLTFOMLMNTNPRRTSKYVEYFPOFKIEKNYEMKQYLALGLKDI 300
 DB 241 LPDDDLSEIESKLSFONIMDMTNSRKMSQYVNVFLPQFKIEKDYEMRSHLSGVLEDF 300
 QY 301 DESKADLSGIASGRLTISRMMHKSYTEEGTEATATGSNIVEKQLPOSTLERADHP 360
 301 DESKADLSGIASGRLTISRMMHKSYTEEGTEATATGSNIVEKQLPOSTLERADHP 360
 DB 301 VESRADLSGIASGRLTISRMMHKSYTEEGTEATATGSNIVEKQLPOSTLERADHP 360
 QY 361 FLEFVIRKDIILFSGKVS 380
 361 FLEFVIRKDIILFSGKVS 380
 DB 361 FLEFVIRKDIILFSGKVS 380

RESULT 9
 AAY08256
 ID AAY08256 standard; Protein; 368 AA.
 XX
 AC AAY08256;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Mouse megin protein.
 XX
 KW Megin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 KM human; rat; murine.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..368
 XX /note= "partial sequence"
 FT
 XX
 PN MO9915652-A1.
 XX
 PD 01-APR-1999.
 XX
 PE 22-SEP-1998; 98MO-JP04269.
 XX
 PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 XX (MIYA/) MIYATA T.
 PI Miyata T;
 XX
 XX
 DR WPI; 1999-276983/23.
 XX
 DR N-PSDB; AAX56714.
 XX
 PT Megin protein expressed specifically in mesangial cells

XX
 PS Claim 1; Page 76-79; 100pp; Japanese.
 XX
 CC This invention describes the isolation of novel megin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as Iga nephropathy.
 XX
 SQ Sequence 368 AA;
 Query Match 74.4%; Score 1458; DB 20; Length 368;
 Best Local Similarity 73.4%; Pred. No. 1e-123;
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;

QY 13 FNLFREKDDNOGNVFFSSLSIFAALALVRLGAODDSLQIDKLHVNTASGIVSSNS 72
 13 FNLFREKDDNOGNVFFSSLSIFAALALVRLGAODDSLQIDKLHVNTASGIVSSNS 72
 DB 1 FDLFREKDDNOGNVFFSSLSIFAALALVRLGAODDCAQIDKLHVNTASGIVSSNS 72
 QY 73 QSGIAGSOLKRVFSDINASHKDYLSIVNGLFAEKVYGFHKDYIECAEKLYDAKVERVDFT 132
 73 QSGIAGSOLKRVFSDINASHKDYLSIVNGLFAEKVYGFHKDYIECAEKLYDAKVERVDFT 132
 DB 61 QPGLQYOLKRVLADINSSHKDYELSTATGVFAEKVYDFHKNYIECAENLYNAKVERVDFT 120
 QY 133 NMLEDTRRNINKVVENETHGKIKNYIGGGISSAVVLYNNAVYFKGKQSAFTKSETIN 192
 133 NMLEDTRRNINKVVENETHGKIKNYIGGGISSAVVLYNNAVYFKGKQSAFTKSETIN 192
 DB 121 NDVQDTRFKINKVVENETHGKIKNYIGGGISSAVVLYNNAVYFKGKQSAFTKSETIN 180
 QY 193 CHFKSPKCSGNAVAMHOERKFNLSVIDPSMKILELRYNGGINMYVLPLENDLSEIENK 252
 193 CHFKSPKCSGNAVAMHOERKFNLSVIDPSMKILELRYNGGINMYVLPLENDLSEIENK 252
 DB 181 CRRSPKPCPKAVNMHOERFNLSVIDPSMKILELRYNGGINMYVLPLENDLSEIENK 240
 QY 253 LTFONLMEWTNPRRTSKYVEYFPOFKIEKNYEMKQYLALGLKDI DESKADLSGIAS 312
 253 LTFONLMEWTNPRRTSKYVEYFPOFKIEKNYEMKQYLALGLKDI DESKADLSGIAS 312
 DB 241 LSFONLMDWTNPRRTSKYVEYFPOFKIEKNYEMKQYLALGLKDI DESKADLSGIAS 300
 QY 313 GGRLYISRMMHKSYTEEGTEATATGSNIVEKQLPOSTLERADHPFLFVIRKDDIIL 372
 313 GGRLYISRMMHKSYTEEGTEATATGSNIVEKQLPOSTLERADHPFLFVIRKDDIIL 372
 DB 301 GGRLYISRMMHKSYTEEGTEATATGSNIVEKQLPOSTLERADHPFLFVIRKDDIIL 360
 QY 373 FSGKVS 380
 373 FSGKVS 380
 DB 361 FSGKVS 380

RESULT 10
 AAB24151
 ID AAB24151 standard; Protein; 368 AA.
 XX
 AC AAB24151;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Mouse megin protein sequence SEQ ID NO:21.
 XX
 KW Megin; mesangium-predominant gene; serpin regulated; nephropathy;
 KM Iga; immunoglobulin A; detection; renal function; renal disorder;
 XX diagnosis; biological sample; blood; urine.
 XX
 OS Mus musculus.
 XX
 PN WO200057189-A1.
 XX
 PD 28-SEP-2000.
 XX
 PE 17-MAR-2000; 2000MO-JP01646.
 XX
 PR 19-MAR-1999; 99JP-0075305.
 XX
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KURO/) KUROKAWA K.
 XX (FUSO/) FUSO PHARM IND LTD.
 PA (MIYA/) MIYATA T.
 XX

PI Miyata T;
 XX
 DR WPI: 2000-611642/58.
 DR N-PSDB: AAC55239.
 XX

PT Evaluating renal function comprises assaying megalin protein in
 PT biological sample
 XX

PS Disclosure: Page 89-91; 93pp; Japanese.
 XX

CC The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megalin protein in biological sample. Also
 CC described are: (1) use of an anti-megalin protein antibody for diagnosing
 CC renal function; and (2) a kit for detecting megalin protein comprising:
 CC (a) anti-megalin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying megalin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the mouse megalin
 CC protein, which is given in the exemplification of the present invention.
 CC
 XX

SQ Sequence 368 AA:

Query Match 74.4%; Score 1458; DB 21; Length 368;
 Best Local Similarity 73.4%; Pred. No. 1e-123;
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;

OY 13 FNLEFREMDDNONGNGVNFSSLSLFAALALVRLGADDSLSQIDKLHVNTASGYGSSNS 72
 DB 1 FDLFREMDSOGNGVNFSSLSIFALTLRLRGAGDCARQIDKALHFNIPSRGSSNN 60

OY 73 QSGLOSOLKRVFSIDINASHKDYLSIVNGLFAEKYGFHKDYIECAEKLVDKVERVDFT 132
 DB 61 QPGLOQLKRVLYADINSSHKDYELSTATGVFAEKYDFHKNYIECAENLYNAKVERVDFT 120

OY 133 NHLEDFRNRINKWVENETHGKIKKNIYEGGSISSAVMLVNAVYFKGKMQSAFTTSETIN 192
 DB 121 NDVQDTRRKINKWENETHGKIKKNIYEGGSISSAVMLVNAVYFKGKMQSAFTTDTLS 180

OY 193 CHEFSPKSGKAVAMHOERKFNLSVIDPSMKILELRYNGGINNYVLLPENDLSEIENK 252
 DB 181 CRFSPPTCPGKAVNMHOBERRFNLSITQOPPMQVLELYHGIGISMYIMLPEDGLCEIESK 240

OY 253 LTFONLMEJTNPRRMTSKYVEVFPPQFIENKNEYEKOYLRAIGLKDIDESSADLSGIAS 312
 DB 241 LSFONLMDTNRKRKKSQYVNVFLPQFIENKNEYETHHLKSLGLKDIDESSADLSGIAS 300

OY 313 GGRLYISRMHMKSYIEVEEGTEATAGSNIVEKQLPOSTLFRADHPFLPIYRKDDITL 372
 DB 301 GGRLYVSKLHMKSFIEVSEEGTEATAENNIVEKQLESTIVFRADRPFLPIYIKKNDITL 360

OY 373 FSGKVSCLP 380
 DB 361 FTGKVSCLP 368

RESULT 11

AAB83077 ID AAB83077 standard; Protein: 368 AA.

AC AAB83077;

XX 10-JUL-2001 (first entry)

XX Murine megalin protein.

XX Mouse; mesangial cell proliferative nephritis; nephrotropic;
 KW transgenic mouse; glomerular disease; animal model; drug screening.
 XX Mus musculus.

OS WO200124628-A1.
 XX
 PN

XX 12-APR-2001.
 PD
 XX 06-OCT-2000; 2000WO-JP06988.
 XX
 PF 06-OCT-1999; 99JP-0285736.
 XX

PA (KUROO) KUROKAWA K.
 PA (MIYATA) MIYATA T.
 XX

PI Miyata T;
 XX

DR WPI: 2001-300136/31.
 DR N-PSDB: AAF82440.
 XX

PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments
 XX

PS Disclosure: Page 52-53; 62pp; Japanese.
 XX

CC The present sequence is murine megalin protein. The human megalin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analysing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 CC
 XX

SQ Sequence 368 AA:

Query Match 74.4%; Score 1458; DB 22; Length 368;
 Best Local Similarity 73.4%; Pred. No. 1e-123;
 Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;

OY 13 FNLEFREMDDNONGNGVNFSSLSLFAALALVRLGADDSLSQIDKLHVNTASGYGSSNS 72
 DB 1 FDLFREMDSOGNGVNFSSLSIFALTLRLRGAGDCARQIDKALHFNIPSRGSSNN 60

OY 73 QSGLOSOLKRVFSIDINASHKDYLSIVNGLFAEKYGFHKDYIECAEKLVDKVERVDFT 132
 DB 61 QPGLOQLKRVLYADINSSHKDYELSTATGVFAEKYDFHKNYIECAENLYNAKVERVDFT 120

OY 133 NHLEDFRNRINKWVENETHGKIKKNIYEGGSISSAVMLVNAVYFKGKMQSAFTTSETIN 192
 DB 121 NDVQDTRRKINKWENETHGKIKKNIYEGGSISSAVMLVNAVYFKGKMQSAFTTDTLS 180

OY 193 CHEFSPKSGKAVAMHOERKFNLSVIDPSMKILELRYNGGINNYVLLPENDLSEIENK 252
 DB 181 CRFSPPTCPGKAVNMHOBERRFNLSITQOPPMQVLELYHGIGISMYIMLPEDGLCEIESK 240

OY 253 LTFONLMEJTNPRRMTSKYVEVFPPQFIENKNEYEKOYLRAIGLKDIDESSADLSGIAS 312
 DB 241 LSFONLMDTNRKRKKSQYVNVFLPQFIENKNEYETHHLKSLGLKDIDESSADLSGIAS 300

OY 313 GGRLYISRMHMKSYIEVEEGTEATAGSNIVEKQLPOSTLFRADHPFLPIYRKDDITL 372
 DB 301 GGRLYVSKLHMKSFIEVSEEGTEATAENNIVEKQLESTIVFRADRPFLPIYIKKNDITL 360

OY 373 FSGKVSCLP 380
 DB 361 FTGKVSCLP 368

RESULT 12

AAV15155 ID AAV15155 standard; Protein: 379 AA.

AC AAV15155;

XX 07-FEB-2000 (first entry)
 XX
 DT
 XX

DE Human prostate serine protease inhibitor (PROSTAPIN).
 XX
 XX PROSTAPIN; serine protease inhibitor; serpin family; prostate; human;
 KM prostate-specific tumor suppressor; apoptosis-inducer; diagnosis;
 KM apoptosis-modulator; prognosis; treatment; prostate cancer;
 KM metastatic prostate cancer.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Inhibitory-site 330..347
 FT /label= "Serine-protease-inhibitory-domain
 FT /note= "Also known as reactive-site loop (RSL) and
 FT contains a hinge region and a variable region. RSL
 FT binds to protease active site during inhibition"
 FT 330..338
 FT Region /label= "RSL-hinge region
 FT /note= "confers stability to the serpin-protease
 FT complex and is highly conserved among the serpins"
 FT 344..345
 FT Cleavage-site /note= "Included in the RSL variable region. PROSTAPIN
 FT undergoes cleavage at this site during inhibition"
 XX
 XX W0958560-A2.
 PN 18-NOV-1999.
 XX
 XX 31-MAR-1999; 99WO-US07123.
 PF
 XX 31-MAR-1998; 98US-0080167.
 PR 15-MAY-1998; 98US-0085720.
 XX
 XX (UROG-) UROGENESYS INC.
 PA (AFAR/) AFAR D E.
 PA (HUBE/) HUBERT R S.
 PA (LEON/) LEONG K.
 PA (RAIT/) RAITANO A B.
 PA (SAFF/) SAFFRAN D C.
 XX
 XX Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
 PI WPI: 2000-039068/03.
 DR N-PSDB; AA229049.
 XX
 XX New isolated prostapin gene, used to develop products for, e.g.
 PT treating and determining susceptibility to prostate cancer
 PT
 PS Claim 1; Fig 1; 63pp; English.
 XX
 XX The present sequence is that of a PROSTAPIN protein which is a member of
 CC serine protease inhibitor (serpin) family and expressed in normal
 CC prostate and locally confined prostate cancer. The sequence is derived
 CC from overlapping sequences of human cDNA clones 11P26 and 103. PROSTAPIN
 CC contains a reactive-site loop characteristic of serpin family in
 CC the carboxy-terminal region. The protein may function as a prostate-
 CC specific tumour suppressor, apoptosis-inducer or apoptosis-modulator.
 CC The sequence is useful in diagnosing, prognosis, monitoring and
 CC treating prostate cancer, particularly advanced stage and metastatic
 CC prostate cancer.
 CC
 XX
 SQ Sequence 379 AA;
 Query Match 40.7%; Score 796.5; DB 21; Length 379;
 Best Local Similarity 42.6%; Pred. No. 1.2e-63;
 Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 114 YIECAEKLYDAKVERVDFTNHLIEDTRRNINKWVENETHGRIKKNVIGEGGISSAVMYLVN 173
 Db 121 YLSCSEKWKYQARLQTVDFEOSTETREKMINAVENKNGKAVMLFGKSTIDPSSVMVLVN 180
 QY 174 AVYFKGKMGSAFLKSEITNHFSPKSGR--AVAMMHQRKNLSVTEDEPSKILELRY 231
 Db 181 TIFYGGRQNKF-----QGNVTVEMVYQIGTFKLAFFKBPQMVLELPY 225
 QY 232 -NGINMYVLLPEN--DLSEIENKLFQNMEMVTPRMYSKYVEVFPQFIKKNVEMK 288
 Db 226 VNNKLSMILLPPIGIANLKOIEKQNGSTHEHTSSNMHREVEVHLPPFKLEIKYELN 285
 QY 289 QYRALGLKIDFESKADLSGIASGGRLYISRMHMSYIEFTEGTEATPAATSNIVEQ 348
 Db 286 SLKPLGVTDLFNQVKADLSGMSPTGLYLKAIHKSYLDVSEGEFPAANAATGDSIAVKS 345
 QY 349 LPOSTLFRADHPFLFVYR--KDIILFSGVSCP 380
 Db 346 LPMRAQFKANHPFLFIRHTHTNTVILFCGLASP 379
 RESULT 13
 AAM47207 standard; Protein; 395 AA.
 ID AAM47207;
 AC AAM47207;
 DT 12-FEB-2002 (first entry)
 XX
 XX Human NOVI protein.
 DE
 XX Human; NOVI; metabolic disorder; neurodegenerative disorder;
 KM immune disorder; haematopoietic disorder; developmental disease; cancer;
 KM retinal disease; feeding disorder; vaccine; infection; gene therapy;
 KM neurological disorder; psychotic disorder; G-protein coupled receptor;
 KM cytoskeletal; antidiabetic; virucide; neuroprotective; noctropic;
 KM analgesic; antidepressant; antimalarial; anticonvulsant; neuroleptic;
 KM antiallergic; antinflammatory; anorectic; antiarthritic;
 KM antiparasitic; antithrombotic; antibacterial; fungicide;
 KM osteopathic; prozoacide; antilucer; hypertensive; hypotensive;
 KM antifertility; vulnery; nephrotoxic; antilipemic; leupin;
 KM chromosome 18.
 XX
 XX Homo sapiens.
 OS
 XX W0200174851-A2.
 PN 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US10039.
 PF
 XX 30-MAR-2000; 2000US-193205P.
 PR 30-MAR-2000; 2000US-193339P.
 PR 05-APR-2000; 2000US-195343P.
 PR 06-APR-2000; 2000US-195005P.
 PR 06-APR-2000; 2000US-195088P.
 PR 10-APR-2000; 2000US-195792P.
 PR 11-APR-2000; 2000US-196556P.
 PR 13-APR-2000; 2000US-197081P.
 PR 14-APR-2000; 2000US-197087P.
 PR 14-APR-2000; 2000US-197525P.
 PR 29-MAR-2001; 2001US-0823187.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Majumder K, Spaderna SK, Taupier RJ, Padigaru M, Burgess CE;
 PI Shimkets RA, Spytek KA, Liu X, Patlurajan M, Gusev VY;
 XX WPI: 2001-626379/72.
 DR N-PSDB; ABA01980.
 XX
 PT New G protein-coupled receptor related polypeptides and polynucleotides

PT for diagnosis, prevention and treatment of metabolic,
PT neurodegenerative, retinal, immune, hematopoietic disorders, diabetes,
PT obesity and infections -
PS Claim 1: Page 9: 194pp: English.

CC The present invention provides the protein and coding sequences of novel
CC human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4,
CC NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can
CC be used in the treatment of NOVA related diseases, including cancer,
CC metabolic, neurodegenerative, immune, haematopoietic, developmental,
CC retinal, feeding, neurological and psychiatric diseases and disorders and
CC infections. The present sequence is the NOV1 protein, the gene for which
CC is found on chromosome 18. The NOV1 protein shares homology with the
CC leupin protein.

XX Sequence 395 AA:

Query Match 39.4%; Score 771.5; DB 22; Length 395;
Best Local Similarity 40.3%; Pred. No. 2,4e-61;
Matches 159; Conservative 89; Mismatches 132; Indels 15; Gaps 5;

OY 1 MASLAANAEPCNLFREDDNOGNGVFPSSLSFPAALALVRLGADDSLSQIDKLHY 60
DB 1 MDSLVTANTKFCFDLFOEIGKDDRHKNIFSPSLSAALGMVRLGARSDSAHQIDEVLHF 60
OY NTASG-----GNSSNSOGLOS-QLKRVFSDINASHKQVDSLVNGLFAEKYGFHKKOY 114
DB 61 NESQNESEKPEAGSLNNESEGLVSCYFGQLSKLDRIKTDYTLSTANRLVGEQEPICQEX 120
OY 115 IECAEKLDAKVERVDFTNHLIEDTRRNINKVENETHGKIKNYIGEGISSAVMYLVNA 174
DB 121 LQGVYFIYHTTIESVDFQKNPEKRSQEIINFVWCOSQGRKDLFSNDATIAEVLYLVNA 180
OY 175 VYFKKMQSAFTKSETINCHFKSPKCGKAVAMHDEKFNLSVIEDPSKILELRY-NG 233
DB 181 YFRKAWETVFDHENTVDAPFCNLNENKSVKMMTKGLRIEIEVKAKQIEMRYTKG 240
OY 234 GIMMYVLLPEND-----LSEIENKLTFOQNLMENTNPRMTSKYVEVFPQFIKKYEM 287
DB 241 KLSMVFALPSHSDNKLKLELEKTIYKVMAMSSSENNSESVLSFPFLEDDYDL 300
OY 288 KOYLRLGLKIDFDESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATGNSNIYER 347
DB 301 NSILQMGIDIDFETRADLTGISPNLTLSTKIHTFEVEVENGQAATAAGAVYSER 360
OY 348 QLPSTLFRADAPPLFYIR--KDDIIFSGKVSQP 380
DB 361 SLRSWVEFNANHPLEFIRHNKTQTLIFYGRVCSF 395

RESULT 14
ABBS3280
ID ABBS3280 standard; Protein: 394 AA.

XX ABB53280;

XX 12-FEB-2002 (first entry)

XX Human polypeptide #20.

XX Human; nootropic; neuroprotective; anticonvulsant; antidepressant;
XX human; tranquiliser; antiarrhythmic; cardiant; antiasthmatic;
XX antiinflammatory; antilipemic; hepatotropic; vitrucide; antidiabetic;
XX nephrotoxic; anorectic; cytosaltic; vaccine; neurological disease;
XX cardiovascular disease; respiratory disease; liver disease;
XX renal disease; skeletal muscle disease; gastrointestinal disease;
XX placental disease; testicular cancer; male fertility; pancreatic disease.

OS Homo sapiens.
XX WO200181363-A1.
PN

PD 01-NOV-2001.
XX 26-APR-2001: 2001WO-US13360.

XX 27-APR-2000: 2000US-19963P.
XX 11-MAY-2000: 2000US-203336P.
XX 25-MAY-2000: 2000US-207087P.
XX 26-MAY-2000: 2000US-207546P.

PA (SMIK) SMITHKLINE BECHAM CORP.
PA (SMIK) SMITHKLINE BECHAM PLC.

PI Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabinick KS;
PI Lal Y, Xie Q;

DR WPI: 2002-041392/05.
DR N-PDB: ABA90345.

PT Novel polypeptides and polynucleotides useful as a vaccine for
PT preventing and treating diseases associated with the polypeptide, e.g.
PT Alzheimer's disease, dyslipidemia, obesity, diabetes, infertility,
PT asthma, amnesia -

PS Claim 1: Page 90-91: 116pp: English.

CC The invention relates to an isolated polypeptide comprising a 277, 480,
CC 583, 581, 628, 424, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,
CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,
CC 784, 252, 593, 472, 607, 781, 640, 686 or 134 amino acid sequence as
CC given in the specification. The polypeptides, modulators of the
CC polypeptides and antibodies against the polypeptides are useful for
CC treating diseases such as neurological and psychiatric diseases
CC including Alzheimer's, paraneoplastic palsy, Huntington's disease,
CC myotonic dystrophy, anorexia and depression, cardiovascular diseases
CC including congestive heart failure, Hodgkin's disease and myocardial
CC infarction; respiratory diseases including asthma, chronic obstructive
CC pulmonary disease, cystic fibrosis and adult respiratory distress
CC syndrome; liver diseases including hypercholesterolemia, cirrhosis,
CC viral and nonviral hepatitis, type II diabetes mellitus, and impaired
CC glucose tolerance; renal disease including renal failure, acute tubular
CC necrosis and glomerulonephritis; skeletal muscle diseases including
CC Ehlerburg's disease, hypoglycaemia and obesity; gastrointestinal
CC diseases including myotonia congenita and intestinal obstruction; lymph
CC diseases including lymphoedema; diseases of testes including testicular cancer,
CC chorioncarcinoma; diseases of testes including testicular cancer,
CC male reproductive diseases including low testosterone and male
CC infertility; and disease of pancreas including diabetic ketoacidosis,
CC type 1 and 2 diabetes and obesity. The present sequence is a
CC polypeptide of the invention.

XX Sequence 394 AA:

Query Match 39.3%; Score 769; DB 23; Length 394;
Best Local Similarity 40.1%; Pred. No. 4e-61;
Matches 158; Conservative 88; Mismatches 134; Indels 14; Gaps 5;

OY 1 MASLAANAEPCNLFREDDNOGNGVFPSSLSFPAALALVRLGADDSLSQIDKLHY 60
DB 1 MDSLVTANTKFCFDLFOEIGKDDRHKNIFSPSLSAALGMVRLGARSDSAHQIDEVLHF 60
OY 61 NTASG-----GNSSNSOGLOS-QLKRVFSDINASHKQVDSLVNGLFAEKYGFHKKOY 115
DB 61 NKTEPFDQAGSLNNESEGLVSCYFGQLSKLDRIKTDYTLSTANRLVGEQEPICQEYL 120
OY 116 ECAEKLDAKVERVDFTNHLIEDTRRNINKVENETHGKIKNYIGEGISSAVMYLVNAV 175
DB 121 DGVOFYHTTIESVDFQKNPEKRSQEIINFVWCOSQGRKDLFSNDATIAEVLYLVNAV 180
OY 176 YFRKMQSAFTKSETINCHFKSPKCGKAVAMHDEKFNLSVIEDPSKILELRY-NGC 234
DB 181 YFRKAWETVFDHENTVDAPFCNLNENKSVKMMTKGLRIEIEVKAKQIEMRYTKG 240
OY 235 INMYVLLPEND-----LSEIENKLTFOQNLMENTNPRMTSKYVEVFPQFIKKYEMK 288

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 67.9418 Seconds
(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-34

Perfect score: 1959
Sequence: 1 MASLAANAEECFLEFREM.....FLFVRKDIILFGKVSQP 380

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1959	100.0	380	2	US-08-472-659-34
2	1959	100.0	380	2	US-08-474-661-34
3	1959	100.0	380	2	US-08-611-977-34
4	752	38.4	390	4	US-08-568-1478-2
5	747	38.1	390	4	US-09-266-910-3
6	745	38.0	390	4	US-09-266-910-4
7	730.5	37.3	375	1	US-08-121-714-8
8	730.5	37.3	375	1	US-08-477-108A-8
9	730.5	37.3	375	5	PCT-US93-08322-8
10	730.5	37.3	375	5	PCT-US93-08322-8
11	722.5	36.9	379	1	US-08-121-714-4
12	722.5	36.9	379	1	US-08-477-108A-4
13	722.5	36.9	379	5	PCT-US93-08322-4
14	722.5	36.9	379	5	PCT-US93-08322-4
15	715.5	36.5	391	4	US-09-123-912-110
16	715.5	36.5	391	4	US-09-643-597-110
17	715	36.5	400	4	US-09-123-912-112
18	715	36.5	400	4	US-09-643-597-112
19	673	34.4	382	1	US-07-768-286B-6
20	673	34.4	382	1	US-08-487-823B-3
21	673	34.4	382	2	US-08-997-040-3
22	673	34.4	382	2	US-09-203-237-3
23	671	34.3	392	1	US-07-768-286B-4
24	664.5	33.9	415	1	US-07-911-531-19
25	664.5	33.9	415	1	US-07-693-636A-19
26	664.5	33.9	415	1	US-07-768-286B-2
27	664.5	33.9	434	1	US-07-679-052A-15

28	662.5	33.8	415	4	US-09-026-408-4	Sequence 4, Appl1
29	662	33.8	374	1	US-08-464-148-2	Sequence 2, Appl1
30	662	33.8	374	1	US-08-385-500-2	Sequence 2, Appl1
31	662	33.8	374	1	US-08-846-784-2	Sequence 2, Appl1
32	660.5	33.7	438	1	US-07-679-052A-17	Sequence 17, Appl1
33	653.5	33.4	405	1	US-08-121-714-7	Sequence 7, Appl1
34	653.5	33.4	405	1	US-08-477-108A-7	Sequence 7, Appl1
35	653.5	33.4	405	2	US-08-477-112-7	Sequence 7, Appl1
36	653.5	33.4	405	5	PCT-US93-08322-7	Sequence 7, Appl1
37	645	32.9	375	1	US-08-121-714-5	Sequence 5, Appl1
38	645	32.9	375	1	US-08-477-108A-5	Sequence 5, Appl1
39	645	32.9	375	2	US-08-477-112-5	Sequence 5, Appl1
40	645	32.9	375	5	PCT-US93-08322-5	Sequence 5, Appl1
41	642	32.8	376	4	US-09-200-965-2	Sequence 2, Appl1
42	623	31.8	376	1	US-08-464-148-4	Sequence 4, Appl1
43	623	31.8	376	1	US-08-385-500-4	Sequence 4, Appl1
44	623	31.8	376	1	US-08-846-784-4	Sequence 4, Appl1
45	592	30.2	465	3	US-08-948-997-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-472-659-34
; Sequence 34, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5831030uh1ro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5831030om1
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm R.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLPAALALVRLGAQDLSIQIDKLHY 60
DB 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLPAALALVRLGAQDLSIQIDKLHY 60
QY 61 NTASGYGNSSNSGSGLOSLKRVFSDINASHKDYDLSYNGLEFAEKYGFHKYICAEK 120
DB 61 NTASGYGNSSNSGSGLOSLKRVFSDINASHKDYDLSYNGLEFAEKYGFHKYICAEK 120
QY 121 LYDAKVERVDFTNHLIEDTRRNINKNWENETHGKIKNVIGSGISSAVMLVNAVYFEGK 180
DB 121 LYDAKVERVDFTNHLIEDTRRNINKNWENETHGKIKNVIGSGISSAVMLVNAVYFEGK 180
QY 181 WQSAFTKSETINCHFKSPKCSGKAVAMHGERKFNLSVIEDPSMKITELRYNGINMYVL 240
DB 181 WQSAFTKSETINCHFKSPKCSGKAVAMHGERKFNLSVIEDPSMKITELRYNGINMYVL 240
QY 241 LPENDLSEIENKLTQNLMTNPRRMTSKYVEVFFPOFKIEKNYEMKQYLRALGLKDF 300
DB 241 LPENDLSEIENKLTQNLMTNPRRMTSKYVEVFFPOFKIEKNYEMKQYLRALGLKDF 300
QY 301 DESKADLSGIASGRLYISMMHKSYTEVEEGTEATATGSNIVEKQLPSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISMMHKSYTEVEEGTEATATGSNIVEKQLPSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGRVSCP 380
DB 361 FLFVIRKDDIILFSGRVSCP 380

RESULT 2

US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ. ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLPAALALVRLGAQDLSIQIDKLHY 60
DB 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLPAALALVRLGAQDLSIQIDKLHY 60
QY 61 NTASGYGNSSNSGSGLOSLKRVFSDINASHKDYDLSYNGLEFAEKYGFHKYICAEK 120
DB 61 NTASGYGNSSNSGSGLOSLKRVFSDINASHKDYDLSYNGLEFAEKYGFHKYICAEK 120
QY 121 LYDAKVERVDFTNHLIEDTRRNINKNWENETHGKIKNVIGSGISSAVMLVNAVYFEGK 180
DB 121 LYDAKVERVDFTNHLIEDTRRNINKNWENETHGKIKNVIGSGISSAVMLVNAVYFEGK 180
QY 181 WQSAFTKSETINCHFKSPKCSGKAVAMHGERKFNLSVIEDPSMKITELRYNGINMYVL 240
DB 181 WQSAFTKSETINCHFKSPKCSGKAVAMHGERKFNLSVIEDPSMKITELRYNGINMYVL 240
QY 241 LPENDLSEIENKLTQNLMTNPRRMTSKYVEVFFPOFKIEKNYEMKQYLRALGLKDF 300
DB 241 LPENDLSEIENKLTQNLMTNPRRMTSKYVEVFFPOFKIEKNYEMKQYLRALGLKDF 300
QY 301 DESKADLSGIASGRLYISMMHKSYTEVEEGTEATATGSNIVEKQLPSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISMMHKSYTEVEEGTEATATGSNIVEKQLPSTLFRADHP 360
QY 361 FLFVIRKDDIILFSGRVSCP 380
DB 361 FLFVIRKDDIILFSGRVSCP 380

RESULT 3

US-08-611-977-34
Sequence 34, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: KURIHARA, Tatsuya
APPLICANT: ISHIDA, No. 5972886uhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404

```

: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/611,977
: FILING DATE: 06-MAR-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/091,028
: FILING DATE: 14-JUL-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 4-212305
: FILING DATE: 17-JUL-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-067339
: FILING DATE: 04-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: McGowan, Malcolm K.
: REGISTRATION NUMBER: 39,300
: REFERENCE/DOCKET NUMBER: 001560-204
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 34:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 380 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-611-977-34

Query Match      100.0%; Score 1959; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 1e-178;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

: Patent No. 5783422
: GENERAL INFORMATION:
: APPLICANT: Sumitani, Yoshinori
: APPLICANT: Kato, Hiroshi
: APPLICANT: Sekiguchi, Kiyoshi
: APPLICANT: Takeda, Katsumichi
: TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
: TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Scully, Scott, Murphy & Presser
: STREET: 400 Garden City Plaza
: CITY: Garden City
: STATE: New York
: COUNTRY: USA
: ZIP: 11530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/568,147B
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 800,952
: FILING DATE: 02-DEC-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: DIGILIO, Frank S.
: REGISTRATION NUMBER: 31,346
: REFERENCE/DOCKET NUMBER: 8425
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 516-742-4343
: TELEFAX: 516-742-4366
: TELEX: 230 901 SANS UR
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 390 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-568-147B-2

Query Match      38.4%; Score 752; DB 1; Length 390;
Best Local Similarity 41.5%; Pred. No. 1.6e-63;
Matches 165; Conservative 81; Mismatches 126; Indels 26; Gaps 10;

```

Db 356 ---PASTNEEFHCNHPLEFIRONKNTNSILFYGRFSSP 390

RESULT 5

US-09-266-910-3

; Sequence 3, Application US/09266910

; Patent No. 6344362

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Use of a recombinant protein as receptor of a

; TITLE OF INVENTION: hepatitis virus

; NUMBER OF SEQUENCES: 8

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/266,910

; FILING DATE:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 390 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-09-266-910-3

Query Match 38.1%; Score 747; DB 4; Length 390;

Best Local Similarity 41.2%; Pred. No. 4,8e-63;

Matches 164; Conservative 81; Mismatches 127; Indels 26; Gaps 10;

QY 1 MASLAANAEEFCNLFREDDNDNGNVPFSSLSLPAALALVRLGAODDSLSQIDKLHV 60
115 IECAEKLVDKAVRYDFTNHLIEDTRRNKWNENHGRKIKNYIGGGISSAVMTLVNA 174
120 LQAIKRFYQTSVDFANAPESPCKINSWESQINERIKNIPGNGISNTTLVNA 179
QY 175 VIFKGMQSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIDPMSKLTLELYNG- 233
180 IYFKGMQEKRFNEDKEEFKFNKNTYKSIQMRQYTSFHFASLEDDVQAKVLEIPYKGR 239
QY 234 GIMMYVLLPE--NDLSEINIKLTFONLMEWTNPRRTSKYVEFFQFKIEKYEKKOYL 291
240 DLSMIVLLPNEIDGLQLEKLTAEKLMETSLQNNRETRVDLHFRFVEESYDLDKYL 299
QY 292 RALGLDIFDESKADLSGIASGRILYISRMHKSYLEVEEGTEATAAT----GSNIYE 346
300 RIMGWMDIFN-GGADLSGMTGSGVLSGLVHLKAFYEVYEEGAAATAAATVAVFGSS--- 355
QY 347 KOLPOST--LFRADHPFLFVIR--KDDILFSGKVSFCP 380
356 ---PASTNEEFHCNHPLEFIRONKNTNSILFYGRFSSP 390

RESULT 6

US-09-266-910-4

; Sequence 4, Application US/09266910

; Patent No. 6344362

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Use of a recombinant protein as receptor of a

; TITLE OF INVENTION: hepatitis virus

NUMBER OF SEQUENCES: 8

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/266,910

FILING DATE:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 390 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-266-910-4

Query Match 38.0%; Score 745; DB 4; Length 390;

Best Local Similarity 41.2%; Pred. No. 7,5e-63;

Matches 164; Conservative 81; Mismatches 127; Indels 26; Gaps 10;

QY 1 MASLAANAEEFCNLFREDDNDNGNVPFSSLSLPAALALVRLGAODDSLSQIDKLHV 60
115 IECAEKLVDKAVRYDFTNHLIEDTRRNKWNENHGRKIKNYIGGGISSAVMTLVNA 174
120 LQAIKRFYQTSVDFANAPESPCKINSWESQINERIKNIPGNGISNTTLVNA 179
QY 175 VIFKGMQSAFTKSETINCHFKSPKCSKAVAMHQRKFNLSVIDPMSKLTLELYNG- 233
180 IYFKGMQEKRFNEDKEEFKFNKNTYKSIQMRQYTSFHFASLEDDVQAKVLEIPYKGR 239
QY 234 GIMMYVLLPE--NDLSEINIKLTFONLMEWTNPRRTSKYVEFFQFKIEKYEKKOYL 291
240 DLSMIVLLPNEIDGLQLEKLTAEKLMETSLQNNRETRVDLHFRFVEESYDLDKYL 299
QY 292 RALGLDIFDESKADLSGIASGRILYISRMHKSYLEVEEGTEATAAT----GSNIYE 346
300 RIMGWMDIFN-GGADLSGMTGSGVLSGLVHLKAFYEVYEEGAAATAAATVAVFGSS--- 355
QY 347 KOLPOST--LFRADHPFLFVIR--KDDILFSGKVSFCP 380
356 ---PASTNEEFHCNHPLEFIRONKNTNSILFYGRFSSP 390

RESULT 7

US-08-121-714-8

; Sequence 8, Application US/08121714

; Patent No. 5470970

; GENERAL INFORMATION:

; APPLICANT: Sager, Ruth

; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH

; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESS: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 MB

; OPERATING SYSTEM: MS-DOS (Version 5.0)

; SOFTWARE: Wordperfect (Version 5.1)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121.714
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 375
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-121-714-8

Query Match 37.3%; Score 730.5; DB 1; Length 375;
Best Local Similarity 37.9%; Pred. No. 1.7e-61;
Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAANAERCFNLPREMDNNGNANVFSSLSLFAALALVRLGAGDDSLSDIKLHV 60
DB 1 MEOLSTANTHFAVDFRALNEDPTGNIFISPLSISSALAMIFLGTGNTAAYSKALYF 60
QY 61 NTASGYGNSNSQSGLSQKRVESDINASHKDYDLISVNGLFAEKYVGFHKDYIECAE 120
DB 61 DTV-----EDHSRFQSLNADINKPCAPYILKLANLYGEKTYNFLADFLASQK 110
QY 121 LYDAKVERVDFTNLEDRRINKVVENETHGKIKNYIGEGCISSAVMLVNAVYERK 180
DB 111 MYGALIASVDFQAPEDARKKEINEMVKGOTEGKIPDLVYKQMVNMKTLVNAIYFKN 170
QY 181 WQAFKSETINCHFKSPKSGKAVAMHOKRKFNLSYIEDPSMKILELRVNG-GINNVY 239
DB 171 WQEFKMEATRDAPFLNKKDTKYKMYOKKPPYNIEDLCRVLELPQGEKLSMTI 230
QY 240 LLPEN-----DLSEIENKLTFFONLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRA 293
DB 231 LLPDIEDESTGLEKIEKQLLEKLRMTKRENLYLAENVVHLPRFKLEESYDLTSHLAR 290
QY 294 LGLKDIDESKADLSGIASGRLYISRMHKSYLEVTEEGTEATAGSNIVEKQLPOST 353
DB 291 LGVODLFNRGRKADLSGMSGARDLFVSKIIHKSFPVDLNEGTAAAGTGTMLAMPPEE 350
QY 354 LFRADHPELFVIRKDDI--ILFSGK 376
DB 351 NFNADHPFIFFIRHNPANILFLGR 375

RESULT 8
US-08-477-108A-8
Sequence 8, Application US/08477108A
Patent No. 5801001
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: 'Zou, Zhigang
APPLICANT: Anisowicz, Anthony
TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.108A
FILING DATE: June 7, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/121.714
FILING DATE: 09/01/93
APPLICATION NUMBER: 07/938,823
FILING DATE: 09/01/92
APPLICATION NUMBER: 07/844,296
FILING DATE: 02/28/92
APPLICATION NUMBER: 07/662,216
FILING DATE: 02/28/91
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06570/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 375
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-477-108A-8

Query Match 37.3%; Score 730.5; DB 1; Length 375;
Best Local Similarity 37.9%; Pred. No. 1.7e-61;
Matches 146; Conservative 85; Mismatches 135; Indels 19; Gaps 4;

QY 1 MASLAANAERCFNLPREMDNNGNANVFSSLSLFAALALVRLGAGDDSLSDIKLHV 60
DB 1 MEOLSTANTHFAVDFRALNEDPTGNIFISPLSISSALAMIFLGTGNTAAYSKALYF 60
QY 61 NTASGYGNSNSQSGLSQKRVESDINASHKDYDLISVNGLFAEKYVGFHKDYIECAE 120
DB 61 DTV-----EDHSRFQSLNADINKPCAPYILKLANLYGEKTYNFLADFLASQK 110
QY 121 LYDAKVERVDFTNLEDRRINKVVENETHGKIKNYIGEGCISSAVMLVNAVYERK 180
DB 111 MYGALIASVDFQAPEDARKKEINEMVKGOTEGKIPDLVYKQMVNMKTLVNAIYFKN 170
QY 181 WQAFKSETINCHFKSPKSGKAVAMHOKRKFNLSYIEDPSMKILELRVNG-GINNVY 239
DB 171 WQEFKMEATRDAPFLNKKDTKYKMYOKKPPYNIEDLCRVLELPQGEKLSMTI 230
QY 240 LLPEN-----DLSEIENKLTFFONLMEWTNPRMTSKYVEVFPQFKIEKNYEMKQYLRA 293
DB 231 LLPDIEDESTGLEKIEKQLLEKLRMTKRENLYLAENVVHLPRFKLEESYDLTSHLAR 290
QY 294 LGLKDIDESKADLSGIASGRLYISRMHKSYLEVTEEGTEATAGSNIVEKQLPOST 353
DB 291 LGVODLFNRGRKADLSGMSGARDLFVSKIIHKSFPVDLNEGTAAAGTGTMLAMPPEE 350
QY 354 LFRADHPELFVIRKDDI--ILFSGK 376
DB 351 NFNADHPFIFFIRHNPANILFLGR 375

RESULT 9

Db 171 WOEFMKEATRDAPFRLNKDKTKYKMMYQKKKFPYNIEDLCKRVLELPHYQKELSMII 230
Qy 240 LLPEN-----DLSEIENKLTFFONLMEWTNPRMTSKYVEVFFPOFKIEKNYEMKOYRA 293
Db 231 LLPDDIEDESGLEKIEQLTKLEWTKPENLTLAEVNLHPFKLEESYDITSLHAR 290
Qy 294 LGLADIPESKADLSGASGRLYISRMHKSXYIEVTEEGTEATGATGSNIVEKOLPOST 353
Db 291 LGVODLFNRKADLSGMSGARDLVYSKIHKSFPVDLNEEGTEAATACTIMLMLPBE 350
Qy 354 LFRADHPFLFYIRKDDI--ILFSGK 376
Db 351 NFADHPFLFYIRHNPANILFLGR 375

RESULT 11
US-08-121-714-4
; Sequence 4, Application US/08121714
; Patent No. 5470970
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,714
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-121-714-4

Query Match 36.9%; Score 722.5; DB 1: Length 379;
Best Local Similarity 38.6%; Pred. No. 1e-60;
Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

Qy 1 MASLAANAEEFCNLFREMDNNGNGNVEFFSLSFALALVRLGAODSLSDIKLHV 60
Db 1 MEQSSANTRALDLFLALSENPNAGNIFISPFSSISAMAVNVLDTGRNTAAQLSKTHF 60
Qy 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSTVNGLFAEKVYGFHHDYIECAEK 120

Db 61 NTV-----EEVHSRFQSLNADINKRCASYLILKLANRLYGEKTYNPLPEFLVSTOK 110
Qy 121 LYDAKVERVDFTNHLIEDTRRINKRYVENETNGIKINAVYEGGSISSAMVLYNAVYFGK 180
Db 111 TYGADLADVDOHASEDARKTINOVVKOTEGKIPELLASQWDMTKLVYVNLVYFGN 170
Qy 181 WSAFTKETJINCHFKSPKCGKAVAMHOERFNLSTYEDPSMKILELRNG-GINNV 239
Db 171 WKDFMKAATINAPFRLNKKDKRYKMMYQKKKFPYNIEDLCKRVLELPHYQKELSMII 230
Qy 240 LLPEN-----DLSEIENKLTFFONLMEWTNPRMTSKYVEVFFPOFKIEKNYEMKOYRA 293
Db 231 LLPDDIEDESGLEKIEQLTKLEWTKPENLTLAEVNLHPFKLEESYDITSLHAR 290
Qy 294 LGLADIPESKADLSGASGRLYISRMHKSXYIEVTEEGTEATGATGSNIVEKOLPOST 353
Db 291 LGVODLFNRKADLSGMSGARDLVYSKIHKSFPVDLNEEGTEAATACTIMLMLPBE 350
Qy 354 LFRADHPFLFYIRKDDI--ILFSGK 380
Db 351 NFADHPFLFYIRHNPANILFLGRFSP 379

RESULT 12
US-08-477-108A-4
; Sequence 4, Application US/08477108A
; Patent No. 5801001
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Zou, Zhigiang
; APPLICANT: Anisowicz, Anthony
; TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
; TITLE OF INVENTION: TUMOR SUPPRESSING ACTIVITY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,108A
; FILING DATE: June 7, 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/121,714
; FILING DATE: 09/01/93
; APPLICATION NUMBER: 07/938,823
; FILING DATE: 09/01/92
; APPLICATION NUMBER: 07/844,296
; FILING DATE: 02/28/92
; APPLICATION NUMBER: 07/662,216
; FILING DATE: 02/28/91
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 06570/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

US-08-477-108A-4

Query Match 36.9%; Score 722.5; DB 1; Length 379;
 Best Local Similarity 38.6%; Pred. No. 1e-60;
 Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

QY 1 MASIAAANAEFCNLPREMDDNGNGVFFSSLSFALALVRLGADDSLSQIDKLLHV 60
 Db 1 MEOLSSANTFALDLFALSENPNAGNIFISPFSSAMAMVFLGRTGNTAOLSKTFHF 60
 QY 61 NTASGYGNSNSQSGLOSLKRVESDINASHKDDLSIVNGLFAEKYGFHKDYIECAEK 120
 Db 61 NTV-----EVSRSQSLNADINKRGASTILKLANRLGERTYNFLPEFLVSTOK 110
 QY 121 LYDAKVERVDFTNHLIEDTRRNINKWVENETHGKIKNYIGEGJSSSAVMVLYNAVYFKGK 180
 Db 111 TYGADLASVDFQHASSEDARTINQWVGQTEGKIPELLASGVDNMTKLVNAVYFKGN 170
 QY 181 WQSAFTSEITNCFKSPKSGKAVAMHQRKFNLSVIDPSMKITELRYNG-GIMMYV 239
 Db 171 WKQKFMKEATVTAPEFLNKDKRTVKMMYOKKFAGYIEDLKCRVLELYOGEELSMVI 230
 QY 240 LPDEN-----DSEIENKLTFLONLMEWTNPRRMTSKYVEFPPOKIEKNYEMKOYLRA 293
 Db 231 LRPDDIEDESTGLKIEQLTEKLEHMTKPEMDLFEVAVSLPRKLEBSYTLNDSPLAR 290
 QY 294 LGLKIDFDESKADLSGIASGRLYISRMHKSYLEVTEEGTEAATGSIIVKOLPOST 353
 Db 291 LGVODLFNSKADLSGMSGARDIFISKIVKSPFEVNEEGTEAATAATGATFCMLMPEE 350
 QY 354 LFRADHPEFLPVIRKDD--ILFSGKYSCP 380
 Db 351 NFTAADHPEFLFIRHNSGSLFLGFRPSP 379

RESULT 13

US-08-477-112-4

Sequence 4, Application US/08477112
 Patent No. 5905023

GENERAL INFORMATION:
 APPLICANT: Sager, Ruth
 TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 555X
 OPERATING SYSTEM: MS-DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,112
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/121,714
 FILING DATE: 09/01/93
 APPLICATION NUMBER: 07/938,823
 FILING DATE: 09/01/92
 APPLICATION NUMBER: 07/844,296
 FILING DATE: 02/28/92
 APPLICATION NUMBER: 07/662,216
 FILING DATE: 02/28/91
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 06570/002003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 379

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-477-112-4

Query Match 36.9%; Score 722.5; DB 2; Length 379;
 Best Local Similarity 38.6%; Pred. No. 1e-60;
 Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

QY 1 MASIAAANAEFCNLPREMDDNGNGVFFSSLSFALALVRLGADDSLSQIDKLLHV 60
 Db 1 MEOLSSANTFALDLFALSENPNAGNIFISPFSSAMAMVFLGRTGNTAOLSKTFHF 60
 QY 61 NTASGYGNSNSQSGLOSLKRVESDINASHKDDLSIVNGLFAEKYGFHKDYIECAEK 120
 Db 61 NTV-----EVSRSQSLNADINKRGASTILKLANRLGERTYNFLPEFLVSTOK 110
 QY 121 LYDAKVERVDFTNHLIEDTRRNINKWVENETHGKIKNYIGEGJSSSAVMVLYNAVYFKGK 180
 Db 111 TYGADLASVDFQHASSEDARTINQWVGQTEGKIPELLASGVDNMTKLVNAVYFKGN 170
 QY 181 WQSAFTSEITNCFKSPKSGKAVAMHQRKFNLSVIDPSMKITELRYNG-GIMMYV 239
 Db 171 WKQKFMKEATVTAPEFLNKDKRTVKMMYOKKFAGYIEDLKCRVLELYOGEELSMVI 230
 QY 240 LPDEN-----DSEIENKLTFLONLMEWTNPRRMTSKYVEFPPOKIEKNYEMKOYLRA 293
 Db 231 LRPDDIEDESTGLKIEQLTEKLEHMTKPEMDLFEVAVSLPRKLEBSYTLNDSPLAR 290
 QY 294 LGLKIDFDESKADLSGIASGRLYISRMHKSYLEVTEEGTEAATGSIIVKOLPOST 353
 Db 291 LGVODLFNSKADLSGMSGARDIFISKIVKSPFEVNEEGTEAATAATGATFCMLMPEE 350
 QY 354 LFRADHPEFLPVIRKDD--ILFSGKYSCP 380
 Db 351 NFTAADHPEFLFIRHNSGSLFLGFRPSP 379

RESULT 14

PCT-US93-08322-4

Sequence 4, Application PC/TUS9308322

GENERAL INFORMATION:
 APPLICANT: Sager, Ruth
 TITLE OF INVENTION: MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 555X
 OPERATING SYSTEM: MS-DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/08322
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/938,823
 FILING DATE: 09/01/92
 APPLICATION NUMBER: 07/844,296
 FILING DATE: 02/28/92

```

1  APPLICATION NUMBER: 07/662,216
2  FILING DATE: 02/28/91
3  ATTORNEY/AGENT INFORMATION:
4  NAME: Fisher, Janis K.
5  REGISTRATION NUMBER: 34,819
6  REFERENCE/DOCKET NUMBER: 00530/072001
7  TELECOMMUNICATION INFORMATION:
8  TELEPHONE: (617) 542-5070
9  TELEFAX: (617) 542-8906
10 TELE: 200154
11 INFORMATION FOR SEQ ID NO: 4:
12 SEQUENCE CHARACTERISTICS:
13     LENGTH: 379
14     TYPE: amino acid
15     STRANDEDNESS:
16     TOPOLOGY: linear
17 PCT-US93-08322-4

```

Query Match	36.9%	Score 722.5:	DB 5,	Length 379;
Best Local Similarity	38.6%	Pred. No. 1e-60:		
Matches 150;	Conservative 74;	Mismatches 146;	Indels 19;	Gaps 4;

```

QY 1 MALLAANAEFCNLEFREDDNOGNCNVEFSSLEFALALVLRGAODDSLOQDLHH 60
Db 1 MEOLSSANTRFALDLEFLALSENPPAGNITFISFSSMAANVFLGTRNTAAOJSKTFHF 60
QY 61 NTASGYGNSSNOSGLOLKRVFSDINASHKDYDLSTVNGLEFAEKYVGHKDYTECAEK 120
Db 61 NTV-----EEVHSRFQSLDINDINKRGASVILKLANRLYGEKTYNFLPELVSTOK 110
QY 121 LYQAKVERDFTNHLIEDTRRNINKKANVENETHGKIKNVYEGEGISSAAMVLVYNAVYFKG 180
Db 111 TYGADLASVDFOJASDARKTINOWKVGOTEKIPELLASGMVNMTRKLYLVNAVYFGN 170
QY 181 MOSAFTKSEFINCHFEKSPKCGAAVAMMHQOEKFLVSTIEDPSMKIILELRNG-GIMMY 239
Db 171 WKQKFEKETTAPFLINKKDKRKTYKMMIOKKKRFYGTIEDCKCVLELPQGBEELSMVI 230
QY 240 LLEPEN-----DLSEIENKLFQONLMEWTNPRKRMYSKVEVEFFQFIENKNEYKQYLRA 293
Db 231 LLEPDIEDSTGLKLEBQLTLEKLEHMETKRENDLFIENVNLSLRFKRESEYTLNSDLAR 290
QY 294 LGLKIDFDESKADLSGIASGRLYISRMHKSYLEVTEEGEATPAATGNSIIVEKOLPOST 353
Db 291 LGVODLFNSSKADLSGMSGARDIFISKIYHKSFVEVNEEGTEEAANAATAGIATPCMLPBE 350
QY 354 LFRADHPLFLIRKDD--ILFSGVYSCP 380
Db 351 NFWADHPLFLIRHNSSGSLFLGRRSSP 379

```

```

RESULT 15
US-09-123-912-110
: Sequence 110, Application US/09123912A
: Patent No. 6312695
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Wang, Tonglong
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
: FILE REFERENCE: 210121.455C1
: CURRENT APPLICATION NUMBER: US/09/123,912A
: CURRENT FILING DATE: 1998-07-27
: PRIOR APPLICATION NUMBER: 09/004,802
: PRIOR FILING DATE: 1998-03-18
: NUMBER OF SEQ. ID NOS: 114
: SOFTWARE: PatentIn Ver. 2.0
: SEQ. ID NO 110
: LENGTH: 391
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-123-912-110

```

```

Best Local Similarity 38.0%; Pred. No. 4.9e-60.
Matches 152; Conservative 78; Mismatches 141; Indels 29; Gaps 7.
QY      1 MASLAANAEECFNLIFREMDNOCGNVFFSSLFLAALALVRLGAODDSLQIDKLHV 60
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      1 MDSLGAVSTRIGFDLFEKL - KKTNDGNIFEFSPVIGLITAGVILGTGCRATASQLEEVH- 58
QY      61 NTASGCGNSSNSQCSLOSQLRVPSDIAASHK-----DYDLSVGNLFAEK 106
      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      59 -----SEKERSSSIKAEKEVEIENTFAVHOFQRLTEISKLNDELNTINLFLFEK 112
QY      107 VYGFHKQVIECAEKLYDQAKVERVDPTNLLEPTNRINIKMVENETHGKIRKIVGEGGSS 166
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      113 TYLFQRLKLDYVERKTYHAASLEVPDYNADSRKKINSWESKITEKINDLEPPDSISS 172
QY      167 AVMLVNVVYFKGQKQSAFTKSETINCHFKSPBKCGKAVAMHQERKFNLSVTEPMSKI 226
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      173 TKLVYVNVVYFKGQMDREFKKEKNTKEEFKFMNKSRSVKQMTQSHFSFTLELDQKI 232
QY      227 LELRY -NGGINVYLLPENDLSEIE---NKLTQNLMEVNTNRRMTSKYVEVFPQFKIE 282
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      233 LGIPKKNNDLSNFWLLP -NDIDGTEKIIDIKIPEKLVEMTSPGCHHEERKVNHLPRFEVE 291
QY      283 KNYEKKQTLRALGLKDIFDESKAADLSGASGRGLYSKMMKSYLEVTEDEGTEAFAAAGS 342
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      292 DSYDLEAVLAAMGMDAFSEHKADYSGMSSGGGLVAORFLHSSFAVATEGTEAAAAAGI 351
QY      343 NIVEKQLFQSTLFRADHPFLFYIR--KDIILFSGKVASCP 380
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      352 GFTVTSAPGHENYVCHHPFLFFIRINESNLSILFFERFSSP 391

```

Search completed: July 11, 2003, 12:02:13
Job time : 68.9418 secs

Query Match 36.5%; Score 715.5; DB 4; Length 391;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 : Search time 111.393 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-34
Perfect score: 1959
Sequence: 1 MASLAAMNAFCNLFREMD.....FLFYIRKDDILIFSGKYSKP 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 11641973 residues
Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_NEM_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCRNEM_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/PCRNUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEM_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1959	100.0	380	9 US-10-091-442-34	Sequence 34, Appl
2	1959	100.0	380	9 US-09-140-719-34	Sequence 34, Appl
3	822	42.0	392	9 US-10-198-070-62	Sequence 62, Appl
4	820	41.9	392	9 US-10-198-070-67	Sequence 67, Appl
5	819	41.8	392	9 US-10-198-070-70	Sequence 70, Appl
6	796.5	40.7	379	9 US-10-012-896-946	Sequence 946, App
7	796.5	40.7	379	9 US-09-895-814-946	Sequence 946, App
8	771.5	39.4	390	9 US-09-823-187-2	Sequence 2, Appl
9	752	38.4	390	9 US-09-823-187-29	Sequence 29, Appl
10	752	38.4	390	9 US-09-823-187-30	Sequence 29, Appl
11	748	38.2	390	9 US-09-823-187-27	Sequence 27, Appl
12	743	37.9	390	9 US-09-823-187-28	Sequence 28, Appl
13	740	37.8	390	9 US-08-731-566-2	Sequence 2, Appl
14	738.5	37.7	617	9 US-10-012-896-947	Sequence 947, App
15	738.5	37.7	617	9 US-08-885-814-947	Sequence 947, App
16	722.5	36.9	437	10 US-09-925-301-1200	Sequence 1200, App
17	715.5	36.5	391	9 US-10-157-031-355	Sequence 355, App
18	715.5	36.5	391	9 US-09-466-396A-110	Sequence 110, App
19	715.5	36.5	391	10 US-09-735-705-110	Sequence 110, App

20	715.5	36.5	391	10 US-09-850-716A-110	Sequence 110, App
21	715.5	36.5	391	10 US-09-897-778-110	Sequence 110, App
22	715	36.5	400	9 US-09-466-396A-112	Sequence 112, App
23	715	36.5	400	10 US-09-735-705-112	Sequence 112, App
24	715	36.5	400	10 US-09-850-716A-112	Sequence 112, App
25	715	36.5	400	10 US-09-897-778-112	Sequence 112, App
26	672.5	34.3	340	10 US-09-764-898-169	Sequence 169, App
27	664.5	33.9	415	9 US-10-017-128-4	Sequence 4, Appl
28	664.5	33.9	423	9 US-10-106-658-4538	Sequence 4538, App
29	662.5	33.8	415	10 US-09-902-684-4	Sequence 4, Appl
30	661	33.7	374	9 US-09-974-298-95	Sequence 95, Appl
31	654	33.4	385	9 US-09-847-208-101	Sequence 101, App
32	652	33.3	409	9 US-10-267-311-55	Sequence 55, Appl
33	652	33.3	948	9 US-10-267-311-21	Sequence 21, Appl
34	642	32.8	459	10 US-09-925-300-1440	Sequence 1440, App
35	608	31.0	315	7 US-08-731-566-4	Sequence 4, Appl
36	600.5	30.7	377	9 US-09-823-187-31	Sequence 31, Appl
37	592	30.2	465	9 US-09-987-021-6	Sequence 6, Appl
38	592	30.2	465	10 US-09-957-485-6	Sequence 6, Appl
39	566	28.9	464	10 US-09-414-834-1	Sequence 27, Appl
40	544	27.8	377	9 US-10-165-605A-27	Sequence 27, Appl
41	544	27.8	377	10 US-09-910-430-27	Sequence 27, Appl
42	498.5	25.4	410	9 US-09-987-021-2	Sequence 2, Appl
43	498.5	25.4	410	10 US-09-957-485-2	Sequence 2, Appl
44	477.5	24.4	217	10 US-09-729-674-134	Sequence 134, App
45	476.5	24.3	402	9 US-10-135-629-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUTSUMOTO, Masafumi
IWASA, Fuyuki
TSURUOKA, No. US20020164711A1uo
NAKAZATO, Hiroshi
MURA, Kenju
ISHIDA, No. US20020164711A1uhfro
KURIHARA, Tatsuya
YAMACHI, Kozo
YAMAGUCHI, No. US20020164711A1omf
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-Aug-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-Jun-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-Jul-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-Jul-1992
APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 380 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 34:
 US-10-091-442-34

Query Match 100.0%; Score 1959; DB 9; Length 380;
 Best Local Similarity 100.0%; Pred. No. 6,8e-161;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAODDSLQIDKLHV 60
 DB 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAODDSLQIDKLHV 60
 QY 61 NTASGYGNSNSGSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 DB 61 NTASGYGNSNSGSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 QY 121 LYDAKERYVDFTNHLDTNRNINKWVENETHGKIKVYIGEGGISSAVMVLVNAVYFKGK 180
 DB 121 LYDAKERYVDFTNHLDTNRNINKWVENETHGKIKVYIGEGGISSAVMVLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFSKPCSGKAVAMMOERKFNLSVIDPMSKILTELRNGINMYVL 240
 DB 181 WQSAFTKSETINCHFSKPCSGKAVAMMOERKFNLSVIDPMSKILTELRNGINMYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTPNRRMTSKYVEVFPPQRIEKNYEMKOYLRLGKIDIF 300
 DB 241 LPENDLSEIENKLTFFONLMEWTPNRRMTSKYVEVFPPQRIEKNYEMKOYLRLGKIDIF 300
 QY 301 DESKADLSGIASGRLYISRMHKSYLETEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGRLYISRMHKSYLETEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVS CP 380
 DB 361 FLFVIRKDDIILFSGKVS CP 380

RESULT 2
 US-09-140-719-34
 Sequence 34, Application US/09140719
 Patent No. US20010026931A1
 GENERAL INFORMATION:
 APPLICANT: TSUJIMOTO, Masafumi
 APPLICANT: IWASA, Fuyuki
 APPLICANT: TSUROOKA, No. US20010026931A1uo
 APPLICANT: NAKAZATO, Hiroshi
 APPLICANT: MIURA, Kenju
 APPLICANT: ISHIDA, No. US20010026931A1uhtto
 APPLICANT: KURIHARA, Tatsuya
 APPLICANT: YAMACHIKI, Kozo
 APPLICANT: YAMAGUCHI, No. US20010026931A1omi
 TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Burns, Doane, Swecker & Mathis
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/140,719
 FILING DATE: 08-AUG-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/474,661
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/091,028
 FILING DATE: 14-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 4-212305
 FILING DATE: 17-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-067339
 FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 380 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-140-719-34

Query Match 100.0%; Score 1959; DB 10; Length 380;
 Best Local Similarity 100.0%; Pred. No. 6,8e-161;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAODDSLQIDKLHV 60
 DB 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAODDSLQIDKLHV 60
 QY 61 NTASGYGNSNSGSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 DB 61 NTASGYGNSNSGSGLOSLKRVFSDINASHKDYDLSYVGLFAEKYGFHKDYIECAEK 120
 QY 121 LYDAKERYVDFTNHLDTNRNINKWVENETHGKIKVYIGEGGISSAVMVLVNAVYFKGK 180
 DB 121 LYDAKERYVDFTNHLDTNRNINKWVENETHGKIKVYIGEGGISSAVMVLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFSKPCSGKAVAMMOERKFNLSVIDPMSKILTELRNGINMYVL 240
 DB 181 WQSAFTKSETINCHFSKPCSGKAVAMMOERKFNLSVIDPMSKILTELRNGINMYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTPNRRMTSKYVEVFPPQRIEKNYEMKOYLRLGKIDIF 300
 DB 241 LPENDLSEIENKLTFFONLMEWTPNRRMTSKYVEVFPPQRIEKNYEMKOYLRLGKIDIF 300
 QY 301 DESKADLSGIASGRLYISRMHKSYLETEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGRLYISRMHKSYLETEEGTEATAGSNIVEKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVS CP 380
 DB 361 FLFVIRKDDIILFSGKVS CP 380

RESULT 3
 US-10-198-070-62
 Sequence 62, Application US/10198070
 Publication No. US20030109437A1


```

: GENERAL INFORMATION:
: APPLICANT: AVERBACK, PAUL
: APPLICANT: GEMMELL, JACK
: TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
: TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 59003.000008
: CURRENT APPLICATION NUMBER: US/10/198, 070
: CURRENT FILING DATE: 2002-07-19
: PRIOR APPLICATION NUMBER: 60/306,161
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/306,150
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/331,477
: PRIOR FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 62
: LENGTH: 392
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-198-070-62

```

```

Query Match      42.0%: Score 822; DB 9; Length 392;
Best Local Similarity 43.1%: Pred. No. 8.5e-63;
Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

```

```

QY 1 MASIAANAEFCNLFREMDNONGVNFSSLSLFAALALVRLGADDSLSQDKLLHV 60
D 1 MGSISTANVEFCIDVFEKLSNNIGDIFSSLSLALSLVLLGAGETAEQLEKVLHF 60
QY 61 -----NTASGYGNSN-SQSG-LOSOLKRVFSDINASHKDYLSIVGLFAEKYGGPHKD 113
D 61 SHYVDSLKPGFKDSQAGRIHSEFGEVSQINODSCTLSIANRLVGTMTMAFHQ 120
QY 114 YIECAERLYDAKVERVDFTHLEDTRRNINKWVENETHGRIKINVGIGISSAVMYLVN 173
D 121 YLSCSEKMYQARLQTVDFEQSTETRTKINAWENKNGKVANLFGKSTIDPSSVAVLVN 180
QY 174 AVYFKGKQSAFTSETINCHFKSPKSGKAVAMHDEKFNLSVIDPSSKILELY-N 232
D 181 TTYFGGROKMFQYREVKSPFQLEKGNVTEMVYDIGTFKLFVKEPQVLELYPVN 240
QY 233 GGIMVYLLPEN--DLSEINKLTFONLMEWTNPRRTSKYVEFFPOFKIEKNYEMKQY 290
D 241 NKLMTILLPVGINLQIEKQNSGFHEHTSSNMREVEVHLRFLKIEITYELNSL 300
QY 291 LRALGLKIDFESKADLSGASGRLYISRMHKSYLEVTEEGTEATAGTSNIVEKOLP 350
D 301 LKPLGVTDLFNOVKADLSGSPKGLYLSKAIHKSYLEVSEGEETAAGDSIAVKSPL 360
QY 351 OSTLFRAHDFPLFYIR--KDDILFSGKVSCLP 380
D 361 MRAOFKANHPLFFETIRHTHTNTILFCGLASP 392

```

RESULT 4

```

: Sequence 67, Application US/10198070
: Publication No. US20030109437A1
: GENERAL INFORMATION:
: APPLICANT: AVERBACK, PAUL
: APPLICANT: GEMMELL, JACK
: TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
: TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 59003.000008
: CURRENT APPLICATION NUMBER: US/10/198, 070
: CURRENT FILING DATE: 2002-07-19
: PRIOR APPLICATION NUMBER: 60/306,161
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/306,150
: PRIOR FILING DATE: 2001-07-19

```

```

: PRIOR APPLICATION NUMBER: 60/331,477
: PRIOR FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 67
: LENGTH: 392
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-198-070-67

```

```

Query Match      41.9%: Score 820; DB 9; Length 392;
Best Local Similarity 43.1%: Pred. No. 1.3e-62;
Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

```

```

QY 1 MASIAANAEFCNLFREMDNONGVNFSSLSLFAALALVRLGADDSLSQDKLLHV 60
D 1 MGSISTANVEFCIDVFEKLSNNIGDIFSSLSLALSLVLLGAGETAEQLEKVLHF 60
QY 61 -----NTASGYGNSN-SQSG-LOSOLKRVFSDINASHKDYLSIVGLFAEKYGGPHKD 113
D 61 SHYVDSLKPGFKDSQAGRIHSEFGEVSQINODSCTLSIANRLVGTMTMAFHQ 120
QY 114 YIECAERLYDAKVERVDFTHLEDTRRNINKWVENETHGRIKINVGIGISSAVMYLVN 173
D 121 YLSCSEKMYQARLQTVDFEQSTETRTKINAWENKNGKVANLFGKSTIDPSSVAVLVN 180
QY 174 AVYFKGKQSAFTSETINCHFKSPKSGKAVAMHDEKFNLSVIDPSSKILELY-N 232
D 181 TTYFGGROKMFQYREVKSPFQLEKGNVTEMVYDIGTFKLFVKEPQVLELYPVN 240
QY 233 GGIMVYLLPEN--DLSEINKLTFONLMEWTNPRRTSKYVEFFPOFKIEKNYEMKQY 290
D 241 NKLMTILLPVGINLQIEKQNSGFHEHTSSNMREVEVHLRFLKIEITYELNSL 300
QY 291 LRALGLKIDFESKADLSGASGRLYISRMHKSYLEVTEEGTEATAGTSNIVEKOLP 350
D 301 LKPLGVTDLFNOVKADLSGSPKGLYLSKAIHKSYLEVSEGEETAAGDSIAVKSPL 360
QY 351 OSTLFRAHDFPLFYIR--KDDILFSGKVSCLP 380
D 361 MRAOFKANHPLFFETIRHTHTNTILFCGLASP 392

```

RESULT 5

```

: Sequence 70, Application US/10198070
: Publication No. US20030109437A1
: GENERAL INFORMATION:
: APPLICANT: AVERBACK, PAUL
: APPLICANT: GEMMELL, JACK
: TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
: TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 59003.000008
: CURRENT APPLICATION NUMBER: US/10/198, 070
: CURRENT FILING DATE: 2002-07-19
: PRIOR APPLICATION NUMBER: 60/306,161
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/306,150
: PRIOR FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 60/331,477
: PRIOR FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 125
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 70
: LENGTH: 392
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-198-070-70

```

```

Query Match      41.8%: Score 819; DB 9; Length 392;
Best Local Similarity 43.1%: Pred. No. 1.5e-62;
Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

```

QY 1 MASLAANAEFCNLFREMDNDNGNGVFFSSLSLFAALALVRLGAODDSLQIDKLHV 60
1 MSLSTANVEFCIDVKEFLNSNNIGNIFFSSLSLALSMVLLGARGETAEOLEVLHF 60
QY 61 -----WTASGYGNSSN-SQSG-LQSOLKRVFSDINASHKDYDLSYNGLEFAEKVYGFHKD 113
Db 61 SHTVDSLKPGFKSPSCQAGRIHSEGVFEFSQINPDNSCTLSIANRLYGTMTAFHQO 120
QY 114 YIECAEKLDAKVERDFNHLDTFRNRINKWVENETHGKIKVIGEGISSAVVAVLN 173
Db 121 YLSCSEKWOARLQOTYDFDQSTEEETRKMINAVENKTNKVANLFGKSTIDPSSVAVLN 180
QY 174 AYVEFKWOSAFKSETINCHFRSPKCSGKAVAMHOERKNLSVIEDPSMKILELRY -N 232
Db 181 IIVFKGROKRFQVRETVKSPFOLSEGNVTEMVYQIGTFKLAPEKPOQVLELRY 240
QY 233 GGIMVYLLPEN--DLSEIENKLTFOIMLMTNPRMTSKYVEVFPQFKIEKNYEMK 230
Db 241 NKLSMILLPVGIANLKQIEKQINSQTFHEWTSNNMEREVEVHLPRFKLEIKYEINSL 300
QY 291 LRALGLKDIDPESKADLSGASGRLYISRMHKSYLEVTEGTEATAAGSNIVEKOLP 350
Db 301 LKRLGTYDLFNQVAKADLSGMSPTKGLYLSKAIHKSIDYSEEGTEAATAAGDSIAVKS 360
QY 351 QSTFLRADHPELFVIR--KDDIILFSGKVS 380
Db 361 MRAQFKNHPELFIRHTHTNTILFCGKLASP 392

RESULT 6

US-10-012-896-946
Sequence 946, Application US/10012896
Publication No. US20020183251A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqiu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Mantanabe, Yoshihiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 946
LENGTH: 379
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-896-946

Query Match

40.7%; Score 796.5; DB 9; Length 379;

Best Local Similarity 42.6%; Pred. No. 1,36-60;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;
QY 1 MASLAANAEFCNLFREMDNDNGNGVFFSSLSLFAALALVRLGAODDSLQIDKLHV 60
1 MSLSTANVEFCIDVKEFLNSNNIGNIFFSSLSLALSMVLLGARGETAEOLEVLHF 60
Db 61 -----WTASGYGNSSN-SQSG-LQSOLKRVFSDINASHKDYDLSYNGLEFAEKVYGFHKD 113
QY 61 SHTVDSLKPGFKSPSCQAGRIHSEGVFEFSQINPDNSCTLSIANRLYGTMTAFHQO 120
Db 61 SHTVDSLKPGFKSPSCQAGRIHSEGVFEFSQINPDNSCTLSIANRLYGTMTAFHQO 120
QY 114 YIECAEKLDAKVERDFNHLDTFRNRINKWVENETHGKIKVIGEGISSAVVAVLN 173
Db 121 YLSCSEKWOARLQOTYDFDQSTEEETRKMINAVENKTNKVANLFGKSTIDPSSVAVLN 180
QY 174 AYVEFKWOSAFKSETINCHFRSPKCSGK--AVAMHOERKNLSVIDPSMKILELRY 231
Db 181 IIVFKGROKRF-----QGNVTEMVYQIGTFKLAPEKPOQVLELRY 225
QY 233 -NGIMVYLLPEN--DLSEIENKLTFOIMLMTNPRMTSKYVEVFPQFKIEKNYEMK 288
Db 226 VNKLSMILLPVGIANLKQIEKQINSQTFHEWTSNNMEREVEVHLPRFKLEIKYEIN 285
QY 289 QYIRALGLKDIDPESKADLSGASGRLYISRMHKSYLEVTEGTEATAAGSNIVEK 348
Db 286 SLKPLGVYDLFNQVAKADLSGMSPTKGLYLSKAIHKSIDYSEEGTEAATAAGDSIAVKS 345
QY 349 LPQSTFLRADHPELFVIR--KDDIILFSGKVS 380
Db 346 LPMRAQFKNHPELFIRHTHTNTILFCGKLASP 379

RESULT 7

US-09-895-814-946
Sequence 946, Application US/09895814
Publication No. US20020193296A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqiu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 946
LENGTH: 379
TYPE: PRT
ORGANISM: Homo sapiens
US-09-895-814-946

Query Match

40.7%; Score 796.5; DB 9; Length 379;

Best Local Similarity 42.6%; Pred. No. 1.3e-60;
Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

```

QY 1 MASLAANAEFCENLREMDNONGNNGVFFSSLSLFAALALVRLGAODDSLQIDKLLHV 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MGLSTANAEFCADYKELSNNGIDIFFSSLSLLYALSLVGLGANGETAQLEKLVHF 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 -----WNASGCGNSN-SQSG-LOSQLRVSDINASHKRDLSIYNGLEFAEKYVGHKD 113
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 SHIVDSLKRGFKSPKSCQAGRIHSEGVFEFSQINOPDSCTSTIANRLYGTITMAFHQ 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 114 YIECAERLYDAKVERVDFTNHLEDRNRINKWVENETHGKIKNVIGEGISSAVMLVN 173
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 YLSCSEKMYQARLQTVDFEQSTETETRMKINAVENKNGKAVLFGKSTIDPSVAVLVN 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 174 AVYFKKMQSAFTKSETINCHFKSPKSGK--AAMMHORKNLSVIEDPSKILELRY 231
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 TIFYKGOROKF-----QCKNVTEVMYQIGTEKFLAVKPEPOKQVLELPY 225
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 232 -NGCINMYVLLPEN--DLSEIENKLTQONLMEVTPRMTSKYVEVFPQFKIEKNYEM 288
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 226 VNNKLSMILLPFGIANLQIEKQINSCTHEMTSSNMMEVEVHLPRFKLEIKELN 285
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 289 QYRALGLADIEDESKADLSGIASGRLYISRMHKSYLEVTEGETAATGNSNIYEKQ 348
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 286 SLTKPLGVTDLFNQVKAADLSGMSPTKGLYLSKAIHKSXYLDVSEGEFAAATGDSIAVKS 345
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 349 LPOSTLFRADHPPLFYIR--KDDIILFSGKVSCP 380
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 346 LPMRAQFKANHPFLPFIIRHTHTTILFCGKLASP 379
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 8 US-09-823-187-2

```

; Sequence 2, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupler, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-2

Query Match 39.4%; Score 771.5; DB 9; Length 395;
Best Local Similarity 40.3%; Pred. No. 2e-58;
Matches 159; Conservative 89; Mismatches 132; Indels 15; Gaps 5;

```

QY 1 MASLAANAEFCENLREMDNONGNNGVFFSSLSLFAALALVRLGAODDSLQIDKLLHV 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MDSLYTANTRKFCEDLFOEIGKDRKKNIFPSPLSLAALGVRLGARSDSAHOIDEVLHF 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 NTASGY-----GNSNSQSLQS-QLRVSDINASHKRDVSLVNLFLFAEKYGFHKDY 114
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 NEFSQNSKEPPAGSLNNSGLVSCYFGQLSKLDRIYDTYTLSTANLYBQEPICQCY 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 115 IECAERLYDAKVERVDFTNHLEDRNRINKWVENETHGKIKNVIGEGISSAVMLVN 174
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 LDGVYQRYHTTIESVDQKNPKRSROINFWECQSGKIKDLFSKAINAEVLYLVNA 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 175 VYFKKMQSAFTYSETINCHFKSPKSCGKAVAMMHORKNLSVIEDPSKILELRY-NG 233
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 VYFKAKMETYFDHENTVDADFCLNQNNENKSYKMTQKGLYRIGFIEEVKAQILEMRYTKG 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 234 GINMYVLLPEN-----LSEIENKLTQONLMEVTPRMTSKYVEVFPQFKIEKNYEM 287
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 KLSMFLVLPSSHKDNLGLLELERTKTYEKKVAMSSSESVLSPFPFTEEDSYDL 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 288 KOYRALGLADIEDESKADLSGIASGRLYISRMHKSYLEVTEGETAATGNSNIYEK 347
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 NSILQMGIDIDIPETRADLTGISPSPNVLKSLIHTTFVEVDNGTQAAATGAVYSER 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 348 QLPOSTLFRADHPPLFYIR--KDDIILFSGKVSCP 380
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 SLMSWVEFNANHPFLPFIIRHKTOTILFYGVCSVP 395
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 9 US-09-823-187-29

```

; Sequence 29, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupler, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081

```



```

; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentl Ver. 2.1
; SEQ ID NO 27
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-27

```

```

Query Match 38.2% Score 748; DB 9; Length 390;
Best Local Similarity 40.6%; Pred. No. 2e-56;
Matches 160; Conservative 84; Mismatches 134; Indels 14; Gaps 8;

```

```

QY 1 MASIAANAEPFCNLFREMDNGNGVNFSSLSLFAALAVRLGADDSLSQDKLHV 60
D 1 MNSLSEANTFEMFDLFOQFRKSKEN-NIFSPISITSLALGMVLGADNAAQISKVLHF 59
QY 61 -----NTASGYGNSSNSQSG-LOSOLKRVFSDINASHKDYDLSTVNGLFAEKYGFHKDY 114
D 60 DQYTEHTTEKATYHYHDSGVNHQOKLLTEFNKSDAVALKANKLFGEKTYQFLQCY 119
QY 115 IECAEKLYDAKVERVDFTNHLSDTRRNINKWVENETHGKIKNYIGEGISSAAVVLVNA 174
D 120 LDAIKKRYQTSVESTDFANAPESRRKINSWESQTEKIKNLFPPDGTIGNDTLLVNA 179
QY 175 VYRKGKQSAFTSETINCIFKSPKCGKAVAMHOBKRNLSVIEDPSKILRLRYNG- 233
D 180 IYRKGQENKFKKENTKEEFWPKNTYKSVQMMROINSFNFALEVDYQAKVLEIPYK 239
QY 234 GINMYVLLPE--NDLSEIENKLTFOIMLEWNTNPRMTSKYVEVEFFPOFKIEKNEMQYL 291
D 240 DLSMIVLLPNEIDGLQKLEKLTAEKLEMTSLQNMHEVCVDLHLPRFKMEESYDKDTL 299
QY 292 RALGLDIFDESRADSLGASGRLYISRMHKSYLEVTEEGTEATATGSNIYKOLPQ 351
D 300 RTMGQNVIFN-GDADLSGTMWSHGLSVKYLHKAFAVEVEGEVAAAATAVAVVEELSSPS 358
QY 352 ST-LFRADHPFLFYIR--KDDIIFSGKVCSP 380
D 359 TNEEFCNHPFLFFIRONKNTSILEYGRFSSP 390

```

```

RESULT 12
US-09-823-187-28
; Sequence 28, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patlurajan, Meera
; APPLICANT: Shlmkels, Richard A
; APPLICANT: Spaderna, Steven K

```

```

; APPLICANT: Spylek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentl Ver. 2.1
; SEQ ID NO 28
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-28

```

```

Query Match 37.9% Score 743; DB 9; Length 390;
Best Local Similarity 40.6%; Pred. No. 5.5e-56;
Matches 159; Conservative 84; Mismatches 135; Indels 14; Gaps 8;

```

```

QY 1 MASIAANAEPFCNLFREMDNGNGVNFSSLSLFAALAVRLGADDSLSQDKLHV 60
D 1 MNSLSEANTFEMFDLFOQFRKSKEN-NIFSPISITSLALGMVLGADNAAQISKVLHF 59
QY 61 -----NTASGYGNSSNSQSG-LOSOLKRVFSDINASHKDYDLSTVNGLFAEKYGFHKDY 114
D 60 DQYTEHTTEKATYHYHDSGVNHQOKLLTEFNKSDAVALKANKLFGEKTYQFLQCY 119
QY 115 IECAEKLYDAKVERVDFTNHLSDTRRNINKWVENETHGKIKNYIGEGISSAAVVLVNA 174
D 120 LDAIKKRYQTSVESTDFANAPESRRKINSWESQTEKIKNLFPPDGTIGNDTLLVNA 179
QY 175 VYRKGKQSAFTSETINCIFKSPKCGKAVAMHOBKRNLSVIEDPSKILRLRYNG- 233
D 180 IYRKGQENKFKKENTKEEFWPKNTYKSVQMMROINSFNFALEVDYQAKVLEIPYK 239
QY 234 GINMYVLLPE--NDLSEIENKLTFOIMLEWNTNPRMTSKYVEVEFFPOFKIEKNEMQYL 291
D 240 DLSMIVLLPNEIDGLQKLEKLTAEKLEMTSLQNMHEVCVDLHLPRFKMEESYDKDTL 299
QY 292 RALGLDIFDESRADSLGASGRLYISRMHKSYLEVTEEGTEATATGSNIYKOLPQ 351
D 300 RTMGQNVIFN-GDADLSGTMWSHGLSVKYLHKAFAVEVEGEVAAAATAVAVVEELSSPS 358
QY 352 ST-LFRADHPFLFYIR--KDDIIFSGKVCSP 380
D 359 TNEEFCNHPFLFFIRONKNTSILEYGRFSSP 390

```

```

RESULT 13
US-08-731-566-2
; Sequence 2, Application US/08731566
; Publication No. US20030072752A1
; GENERAL INFORMATION:
; APPLICANT: Toshihiko Hibino et al.

```

```

: TITLE OF INVENTION: TWO NOVEL GENES FROM PSORIASIS EPIDERMIS:
: TITLE OF INVENTION: PSORIASIS TYPE I AND PSORIASIS TYPE II
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1775
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/731,566
: FILING DATE: 16-OCT-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,755
: FILING DATE: 20-MAR-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/005,679
: FILING DATE: 17-OCT-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Louis Meyers
: REGISTRATION NUMBER: 35,965
: REFERENCE/DOCKET NUMBER: MGP-040CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 390 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-731-566-2

```

Query Match 37.8%; Score 740; DB 7; Length 390;

Best Local Similarity 41.0%; Pred. No. 1e-55; Mismatches 128; Indels 26; Gaps 10;

```

Matches 103; Conservative 81; Mismatches 128; Indels 26; Gaps 10;
QY 1 MASLAANAEPFCNLEPREDDNGNGVFFSSLSLEPAALALVRLGAQDSSLQIDKLLHY 60
DB 1 MMSLSAANTKPMEDLQFRKSKEN-NITYSPISLSALGMVLLGAKGTAQGIKKVLFH 59
QY 61 -----NTASGYGNSNSQSG-LOSQLRVPSDINASHKDYDLSTVNGLEPAEKVYGFHKDY 114
DB 60 DQVTEYTKRAATYHYDRSGDVHHQFOKLTTEFNKSTDAVELKIAKRLGEEKYELFLOEY 119
QY 115 ISCAEKLVDKAVYEDFTNHLDEPTRNINKWVENHKGKIKVYIGSGSSSAVMVLVA 174
DB 120 LDAIKKFTYTSVESVDFAANAPESRKIKINWESQTNKIKNLPENGINSTTLVLA 179
QY 175 VYFKGKQSAFTKSETINCHFSKPCSGKAVAMHOREFNLSVIDPSKILRLRYNG- 233
DB 180 IYFKGMEKKFKNEKDEKFKPMNKTYKSIQMMKQYTSFHFASLSDYQAKVLEIYKGR 239
QY 234 GINMYVLLPE--NDLSEINLKLTPONLMETNRRRTSKYVEVFPQFIENKYEKOYL 291
DB 240 DLSMIVLLPNEIDGLQRLLEEELTAELKLMEMTSLQNRRRETRVDLHLRFKVEESYDLKDTL 299
QY 292 RALGLKIDFDESKADISLASGRLYISRMHKSYLEVTEEGEATAAAT-----GSNIYE 346
DB 300 RTMGAVDITN-GDADISGMTSGRLVSLVLAHAFVEVEEGEAALAAVVGFGSS--- 355
QY 347 KOLPOST--LFRADHFLFVIR--KDDIILFSGVKSCP 380
DB 356 ----PSTNEFHCHNHFLEFFIRONKNSILFYGRESPP 390

```

RESULT 14

```

US-10-012-896-947
: Sequence 947, Application US/10012896
: Publication No. US20020183251A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yugu
: APPLICANT: Kalos, Michael D.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stolk, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darrick
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aijun
: APPLICANT: Skelky, Yasir A.W.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Hural, John
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Vinals de Bassols, Carlota
: APPLICANT: Foy, Teresa
: APPLICANT: Fanger, Gary R.
: APPLICANT: Mantabe, Yoshihiro
: APPLICANT: Weather, Madeline Joy
: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
: FILE REFERENCE: 210121.427C27
: CURRENT APPLICATION NUMBER: US/10/012,896
: CURRENT FILING DATE: 2001-12-10
: NUMBER OF SEQ ID NOS: 1011
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 947
: LENGTH: 617
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: 90, 215, 225, 232, 233, 289, 309, 318, 322, 331, 332, 367,
: LOCATION: 369, 374, 378, 425, 430
: OTHER INFORMATION: Xaa = Any Amino Acid
: US-10-012-896-947

```

Query Match 37.7%; Score 738.5; DB 9; Length 617;

Best Local Similarity 28.8%; Pred. No. 2.6e-55; Mismatches 121; Indels 237; Gaps 11;

```

Matches 178; Conservative 81; Mismatches 121; Indels 237; Gaps 11;
QY 1 MASLAANAEPFCNLEPREDDNGNGVFFSSLSLEPAALALVRLGAQDSSLQIDKLLHY 60
DB 1 MMSLSAANAEPCLDVFKELNSNIGNIFSSLSLYALSMLGAKGETAEQLEKVLHF 60
QY 61 -----NTASGYGNSNSQSG-LOSQLRVPSDINASHKDYDLSTVNGLEPAEKVYGFHKD 113
DB 61 SHYVDSLKRGFNDSPKCSAGRIHSEGYXFSQINPDNCGLTSLNRLYGRKTAFAHQ 120
QY 114 YIECAEKLVDKAVYEDFTNHLDEPTRNINKWVENHKGKIKVYIGSGSSSAVMVLVA 173
DB 121 YLSCSEKQVQARLQYTFDEQSTETRTJINAWENTKGNVANLFGKSTIDPSSVMVLVA 180
QY 174 AYFKGKQSAFTKSTIN----- 192
DB 181 AIFYKQWONKQFQVETVAVSPQLSEVSLFSDSXOMLEDYTIKQFPRKMXFSBNIGL 240
QY 193 ----- 192
DB 241 GFCWFLLYLFQIFIFPLSDNNFTYHRAFNWRLGLRFGSGRGNPFESXSLGLEFPYTL 300
QY 193 -----CHP-----KSPK----- 199
DB 301 WLCSPAAHXGYLCYFFFXRVXSXGKIKKMXIXYIILPTKIMLAKNPDEVFGRPSTLYTL 360

```

```

Qy 200 -----CS----- 201
Db 361 LEQSLKXPLLNKNGXPLOREVIYNLCSEFTFHAERVEMOISVLKRVISTHICALTY 420
Qy 202 -----GK--AVAMHDERKFNLSVIEDPSMKILELRY-NGGINMYVLLPEN--D 245
Db 421 VSILXSPSSXOGKNTVEMMYQIGTFKLAFAVKEPOMQVLELPYNNKLSMILLPVGIAN 480
Qy 246 LSEIENKLTFOINLMEINPRMRTSKYVEVFPQRIEKNYEMKOYLALGLKIDFDESKA 305
Db 481 LKQIEKQINSGTFHEWTSSSNMMEVEVHLPRKLETKYELNSLKLGYTDLEFNQYKA 540
Qy 306 DLGSIASGRLYISRMHKSXYIEVTEEGTEATAGSIVYKOLPOSTLFRADHPFLFVI 365
Db 541 DLGSMSPYKGLYLKAIHKSXYLDVSEEGTEAATAAGDSIAVKSILPMRAQFANHPFLFPI 600
Qy 366 R-KDDIILFSGKVSXP 380
Db 601 RHTHTNTILFCGKLASP 617

```

```

RESULT 15
US-09-895-814-947
: Sequence 947, Application US/09895814
: Publication No. US20020193296A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jlangchun
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitchem, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jlang, Yugu
: APPLICANT: Kelos, Michael D.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stoik, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darick
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aljun
: APPLICANT: Skeiky, Yasir A.W.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Hural, John
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Vinals de Bassols, Carlota
: APPLICANT: Foy, Teresa
: APPLICANT: Fanger, Gary R.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.427C26
: CURRENT APPLICATION NUMBER: US/09/895,814
: NUMBER OF SEQ ID NOS: 990
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 947
: LENGTH: 617
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: variant
: LOCATION: (1)...(617)
: OTHER INFORMATION: Xaa - Any amino acid
US-09-895-814-947

```

```

Query Match 37.7%; Score 738.5; DB 9; Length 617;
Best Local Similarity 28.8%; Pred. No. 2.6e-55;
Matches 178; Conservative 81; Mismatches 121; Indels 237; Gaps 11;

```

```

Qy 1 MASLAANAFCFNLFRMDNCGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
Db 1 MGSISTANVERCLDVFKELNSNIGDNIFSSLSLVALSVLGLGARETAOLEKVLHF 60

```

```

Qy 61 -----NTASGYGNSSN--SOSG-LOSQLRVPSDINASHKDYDLSTVNGLFPAEKYGFPHKD 113
Db 61 SHYVDSLKPRGRDSCQACRHISEFVYXLSQIOPSSCTLSIANLITGKTMAFHOQ 120
Qy 114 YIECAEKLYDAKVERVDFTNHLDRNINKWVENETGKIKNYIGEGISSAVMYLVN 173
Db 121 YLSCSEKMYQARLQVDFDEQSTEDTRKTINMVEKKTNGKVANLEFGKSTIDPSSVWLVN 180
Qy 174 AVYERKKQSAFTYSETIN----- 192
Db 181 AIYFEGQMNRFQVARETVKSPFOLSEVSIIFSXSOMLEDYIIIXGPRKMXFSEENIGL 240
Qy 193 ----- 192
Db 241 GFCWFELLYLQIFLIFPLSDNNFYHRAPNRBLGILRSGNGENPFESXRSLGLFFPYIL 300
Qy 193 -----CHP-----KSPK----- 199
Db 301 WLCSPAHHXGYLCYFEFFXRVSGKIKKMXIXXYILFLPTKIMLAKNPDFVGRPSYLYIL 360
Qy 200 -----CS----- 201
Db 361 LEQSLKXPLLNKNGXPLOREVIYNLCSEFTFHAERVEMOISVLKRVISTHICALTY 420
Qy 202 -----GK--AVAMHDERKFNLSVIEDPSMKILELRY-NGGINMYVLLPEN--D 245
Db 421 VSILXSPSSXOGKNTVEMMYQIGTFKLAFAVKEPOMQVLELPYNNKLSMILLPVGIAN 480
Qy 246 LSEIENKLTFOINLMEINPRMRTSKYVEVFPQRIEKNYEMKOYLALGLKIDFDESKA 305
Db 481 LKQIEKQINSGTFHEWTSSSNMMEVEVHLPRKLETKYELNSLKLGYTDLEFNQYKA 540
Qy 306 DLGSIASGRLYISRMHKSXYIEVTEEGTEATAGSIVYKOLPOSTLFRADHPFLFVI 365
Db 541 DLGSMSPYKGLYLKAIHKSXYLDVSEEGTEAATAAGDSIAVKSILPMRAQFANHPFLFPI 600
Qy 366 R-KDDIILFSGKVSXP 380
Db 601 RHTHTNTILFCGKLASP 617

```

Search completed: July 11, 2003, 12:37:36
Job time : 112.393 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 76.632 Seconds
(Without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-34

Perfect score: 1959

Sequence: 1 MASLAANAEFCNLFREMD.....FLFVIRKDDILFSGKVSVC 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database:

1: PIR:73:*\n2: PIR:*\n3: PIR:*\n4: PIR:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	752	38.4	390	2	I38201
2	743	37.9	390	2	I38202
3	739.5	37.7	379	2	A42421
4	722.5	36.9	379	2	S27383
5	715.5	36.5	391	2	JC7118
6	711	36.3	388	1	DYCH
7	699	35.7	378	2	S38962
8	664.5	33.9	415	2	A32853
9	662	33.8	374	2	A59273
10	659	33.6	386	1	OACH
11	658.5	33.6	383	2	S11433
12	658.5	33.6	397	2	I39184
13	649	33.1	378	2	A57488
14	648	33.1	376	1	A48681
15	630.5	32.2	415	2	S20047
16	623	31.8	376	2	B59273
17	622	31.8	416	2	S19896
18	592	30.2	465	2	I59611
19	590.5	30.1	375	2	A36888
20	578	29.5	431	1	JX0364
21	567	28.9	433	1	A61435
22	566	28.9	464	1	XKHU3
23	557	28.4	465	1	S28219
24	478.5	24.4	417	2	S19724
25	476.5	24.3	410	2	S70647
26	476	24.3	418	2	S23675
27	468.5	23.9	359	2	D88940
28	467	23.8	410	2	A45457
29	464.5	23.7	418	2	A53120

30	463.5	23.7	433	1	ITHC	alpha-1-antichymot
31	462.5	23.6	418	2	JX0129	contrapsin precurs
32	461.5	23.6	397	2	T06183	serpin - barley
33	460	23.5	418	1	S31507	serine proteinase
34	450	23.0	408	2	S11320	serine proteinase
35	446.5	22.8	405	2	A39088	alpha-1-antiprotei
36	446.5	22.8	406	2	A39339	protein C inhibito
37	444	22.7	366	2	T16119	hypothetical prote
38	438.5	22.4	413	2	JX0267	alpha-1-antiprotei
39	437	22.3	418	2	JH0494	alpha-1-antichymot
40	435.5	22.2	403	2	S08102	serine proteinase
41	434	22.2	416	2	B29131	kallikrein-binding
42	433.5	22.1	413	2	JX0154	alpha-1-antiprotei
43	433	22.1	232	1	DXCH	ovalbumin-related
44	432	22.1	479	1	S41066	heparin cofactor I
45	430	21.9	366	2	T25504	hypothetical prote

ALIGNMENTS

RESULT 1

I38201 Squamous cell carcinoma antigen 1 - human

C:Species: Homo sapiens (man)

C>Date: 23-Feb-1996 #sequence, revision 23-Feb-1996 #text, change 26-May-2000

C:Accession: I38201, I38200, G01631, J70966, J70967

R:Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter, S.D.; Hul, proc. Natl. Acad. Sci. U.S.A. 92, 3147-3151, 1995

A:Title: A serine proteinase inhibitor locus at 18q21.3 contains a tandem duplication

A:Reference number: I38200; MUID:95241462; PMID:7724531

A:Accession: I38201

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-390 <SCCH1>

A:Cross-references: EMBL:U19568; GB:U19558; NID:g1172085

A:Accession: I38200

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-117 <SCH2>

A:Cross-references: EMBL:U19562; NID:g897835; PIDN:AAA6316.1; PID:g897844

R:Silverman, G.A.

A:Reference number: G07968

A:Accession: G01631

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-350, 'G', 352-390 <SIL>

A:Cross-references: EMBL:U19556; NID:g1276435; PID:g1052869

R:Sumitani, Y.; Kishi, F.; Sekiguchi, K.; Kato, H.

A:Title: Squamous cell carcinoma antigen is a new member of the serine protease inhib

A:Reference number: J70966; MUID:92068241; PMID:1958219

A:Accession: J70966

A:Molecule type: mRNA

A:Residues: 1-350, 'G', 352-356, 'A', 358-390 <SUM1>

A:Cross-references: GB:566896; NID:g239551; PIDN:ABB20405.1; PID:g239552

A:Accession: J70967

A:Molecule type: protein

C:Comment: This antigen probably acts as a proteinase inhibitor to modulate the host

C:Genetics:

A:Gene: GDB:SCCA1; SCC

A:Cross-references: GDB:625364; OMIM:600517

A:Map position: 18q21.3-18q21.3

A:Introns: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3

C:Superfamily: antithrombin III

C:Keywords: cysteine proteinase inhibitor; glycoprotein

F:65/93,171,376/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:354/Inhibitory site: Ser (cathepsin L) #status predicted

Query Match 38.4%; Score 752; DB 2; Length 390;
Best Local Similarity 41.5%; Pred. No. 1.3e-45;

Matches 155; Conservative 81; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASLAANAEFCNLFREMDNNGNGNVFSSLSLFAALALVRLGAQDSSLQIDKLLHV 60
 1 MNSLSEANTKFMFDLQOFRKSKEN-NIFESPISITSALGMVLLGAKDNTAOIKVLFH 59
 Db 60 DQYENTTEKATYHYHDSRGVNHVHOFKLTTEFNKSTDAVELKIANKLTGEKTYLQLOEY 119
 QY 61 -----NTASGYGSSNSOG-LOSQLRVFSIDINASHKDYDLSYVGLFAEKVYGFHKDY 114
 115 IECAERLYDAKVERVDFTNHLDEPTNRINKWVENHGRKKNVIGSGGSSSAVMVLVA 174
 Db 120 LDAIKKFYQTSVESYFAPNAPESSRKINSWESQYNEKKNLIPGNGSNTTLVLVA 179
 QY 175 VYFKGWSAFTKSETINCHFKSPKCSGAVAMHBERKFNLSVIDPSMKILELRYNG- 233
 180 IYFKGWEKKFNEDRKEEFKFNKNTYKSIQMMROYTSFHFASLSDVQAKVLEIYYKK 239
 Db 234 GIMMYVLPE--NDLSEIENKLTFOULMEWTPRRRTSKYVEVFPQFKIEKNYKQYL 291
 240 DLSMIVLLNEIDGLQLEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAE 299
 QY 292 RALGLKIDFESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATATATATATATAT 346
 300 RTMGWYDIFN-GDADLSGWTSHGLSVKRLKRAFEVTEEGEAAATAVAVFELSSPS 355
 Db 347 KQLPOST--LFRADHPFLFVIR--KDDIILFSGKXSCP 380
 356 ---PTSTNEEFCHNHPFLFIRQNTKNSILFYGRFSSP 390

RESULT 2

leupin precursor - human
 N:Alternate names: proelaine inhibitor 11 (P11); squamous cell carcinoma antigen 2
 C:Species: Homo sapiens (man)
 C>Date: 23-Feb-1996 #sequence, revision 23-Feb-1996 #text_change 26-May-2000
 C:Accession: I38202; S66675; S57522
 R:Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter, S.D.; Hui, S.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3147-3151, 1995
 A:Title: A serine proteinase inhibitor locus at 18q21.3 contains a tandem duplication of
 A:Reference number: I38200; M0ID:95241462; PMID:7724531
 A:Accession: I38202
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-390 <SCH>
 A:Cross-references: EMBL:U19576; GB:U19569; NID:9852466
 R:Barnes, R.C.; Morrall, D.M.
 FEBS Lett. 373, 61-65, 1995
 A:Title: Identification of a novel human serpin gene: cloning sequencing and expression
 A:Reference number: S66675; M0ID:96013887; PMID:7589435
 A:Accession: S66675
 A:Molecule type: mRNA
 A:Residues: 7-351, 'V', 353-384 <BAR>
 A:Cross-references: EMBL:X89015; NID:9887464; PIDN:CAA61420.1; PID:9887465
 C:Genetics: GDB:SCA2; P11
 A:Gene: GDB:SCA2; P11
 A:Cross-references: GDB:636556; OMIM:600518
 A:Map position: 18q21.3-18q21.3
 A:Introns: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3
 C:Superfamily: antithrombin III
 C:Keywords: glycoprotein; serine proteinase inhibitor
 F:1-24/Domain: signal sequence #status predicted <STC>
 F:23-384/Product: leupin #status predicted <MAT>
 F:65,93,170,376/Binding site: carbonyl (Asn) (covalent) #status predicted
 F:354/Inhibitory site: Leu (unidentified proteinase) #status predicted

Query Match 37.9%; Score 743; DB 2; Length 390;
 Best Local Similarity 40.6%; Pred. No. 5,4e-45;
 Matches 159; Conservative 84; Mismatches 135; Indels 14; Gaps 8;

QY 1 MASLAANAEFCNLFREMDNNGNGNVFSSLSLFAALALVRLGAQDSSLQIDKLLHV 60
 1 MNSLSEANTKFMFDLQOFRKSKEN-NIFESPISITSALGMVLLGAKDNTAOIKVLFH 59
 Db 60 DQYENTTEKATYHYHDSRGVNHVHOFKLTTEFNKSTDAVELKIANKLTGEKTYLQLOEY 119
 QY 61 -----NTASGYGSSNSOG-LOSQLRVFSIDINASHKDYDLSYVGLFAEKVYGFHKDY 114
 115 IECAERLYDAKVERVDFTNHLDEPTNRINKWVENHGRKKNVIGSGGSSSAVMVLVA 174
 Db 120 LDAIKKFYQTSVESYFAPNAPESSRKINSWESQYNEKKNLIPGNGSNTTLVLVA 179
 QY 175 VYFKGWSAFTKSETINCHFKSPKCSGAVAMHBERKFNLSVIDPSMKILELRYNG- 233
 180 IYFKGWEKKFNEDRKEEFKFNKNTYKSIQMMROYTSFHFASLSDVQAKVLEIYYKK 239
 Db 234 GIMMYVLPE--NDLSEIENKLTFOULMEWTPRRRTSKYVEVFPQFKIEKNYKQYL 291
 240 DLSMIVLLNEIDGLQLEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAE 299
 QY 292 RALGLKIDFESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATATATATATATATAT 346
 300 RTMGWYDIFN-GDADLSGWTSHGLSVKRLKRAFEVTEEGEAAATAVAVFELSSPS 355
 Db 347 KQLPOST--LFRADHPFLFVIR--KDDIILFSGKXSCP 380
 356 ---PTSTNEEFCHNHPFLFIRQNTKNSILFYGRFSSP 390

Db 1 MNSLSEANTKFMFDLQOFRKSKEN-NIFESPISITSALGMVLLGAKDNTAOIKVLFH 59
 QY 61 -----NTASGYGSSNSOG-LOSQLRVFSIDINASHKDYDLSYVGLFAEKVYGFHKDY 114
 115 IECAERLYDAKVERVDFTNHLDEPTNRINKWVENHGRKKNVIGSGGSSSAVMVLVA 174
 Db 120 LDAIKKFYQTSVESYFAPNAPESSRKINSWESQYNEKKNLIPGNGSNTTLVLVA 179
 QY 175 VYFKGWSAFTKSETINCHFKSPKCSGAVAMHBERKFNLSVIDPSMKILELRYNG- 233
 180 IYFKGWEKKFNEDRKEEFKFNKNTYKSIQMMROYTSFHFASLSDVQAKVLEIYYKK 239
 Db 234 GIMMYVLPE--NDLSEIENKLTFOULMEWTPRRRTSKYVEVFPQFKIEKNYKQYL 291
 240 DLSMIVLLNEIDGLQLEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAE 299
 QY 292 RALGLKIDFESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATATATATATATATAT 346
 300 RTMGWYDIFN-GDADLSGWTSHGLSVKRLKRAFEVTEEGEAAATAVAVFELSSPS 355
 Db 347 KQLPOST--LFRADHPFLFVIR--KDDIILFSGKXSCP 380
 356 ---PTSTNEEFCHNHPFLFIRQNTKNSILFYGRFSSP 390

RESULT 3

leukocyte elastase inhibitor - horse
 N:Alternate names: plasminogen activator inhibitor-2 homolog
 C:Species: Equus caballus (domestic horse)
 C>Date: 04-Mar-1993 #sequence, revision 18-Nov-1994 #text_change 02-Jun-2000
 C:Accession: A42421; A37276; S34062
 R:Dubin, A.; Travis, J.; Enghild, J.J.; Potempa, J.
 J. Biol. Chem. 267, 6576-6583, 1992
 A:Title: Equine leukocyte elastase inhibitor. Primary structure and identification as
 A:Reference number: A42421; M0ID:92202200; PMID:1551869
 A:Accession: A42421
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-379 <DUB>
 A:Cross-references: PIDN:AMB21885.1; PID:9247842
 A:Experimental source: leukocyte
 A:Note: Sequence extracted from NCBI backbone (NCBIP:89849)
 R:Dubin, A.; Travis, J.; Enghild, J.J.; Potempa, J.
 submitted to the Protein Sequence Database, December 1991
 A:Reference number: A37276
 A:Accession: A37276
 A:Molecule type: protein
 A:Residues: 1-41, 'E', 43-325, 'VD', 326-379 <DUB>
 R:Kordula, T.; Dubin, A.; Schollnick, H.; Koj, A.; Heinrich, P.C.; Rose-John, S.
 Biochem. J. 293, 187-193, 1993
 A:Title: Molecular cloning and expression of an intracellular serpin: an elastase inh
 A:Reference number: S34062; M0ID:93319507; PMID:7687128
 A:Accession: S34062
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-379 <KOR>
 A:Cross-references: GB:M91161; NID:9164240; PIDN:AAA97513.1; PID:9164241
 C:Superfamily: antithrombin III
 C:Keywords: serine proteinase inhibitor

Query Match 37.7%; Score 739.5; DB 2; Length 379;
 Best Local Similarity 38.0%; Pred. No. 9,2e-45;
 Matches 148; Conservative 85; Mismatches 137; Indels 19; Gaps 4;

QY 1 MASLAANAEFCNLFREMDNNGNGNVFSSLSLFAALALVRLGAQDSSLQIDKLLHV 60
 1 MNSLSEANTKFMFDLQOFRKSKEN-NIFESPISITSALGMVLLGAKDNTAOIKVLFH 59
 Db 60 DQYENTTEKATYHYHDSRGVNHVHOFKLTTEFNKSTDAVELKIANKLTGEKTYLQLOEY 119
 QY 61 -----NTASGYGSSNSOG-LOSQLRVFSIDINASHKDYDLSYVGLFAEKVYGFHKDY 114
 115 IECAERLYDAKVERVDFTNHLDEPTNRINKWVENHGRKKNVIGSGGSSSAVMVLVA 174
 Db 120 LDAIKKFYQTSVESYFAPNAPESSRKINSWESQYNEKKNLIPGNGSNTTLVLVA 179
 QY 175 VYFKGWSAFTKSETINCHFKSPKCSGAVAMHBERKFNLSVIDPSMKILELRYNG- 233
 180 IYFKGWEKKFNEDRKEEFKFNKNTYKSIQMMROYTSFHFASLSDVQAKVLEIYYKK 239
 Db 234 GIMMYVLPE--NDLSEIENKLTFOULMEWTPRRRTSKYVEVFPQFKIEKNYKQYL 291
 240 DLSMIVLLNEIDGLQLEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAEKLTAE 299
 QY 292 RALGLKIDFESKADLSGIASGRLYISRMHKSYLEVTEEGTEATATATATATATATATAT 346
 300 RTMGWYDIFN-GDADLSGWTSHGLSVKRLKRAFEVTEEGEAAATAVAVFELSSPS 355
 Db 347 KQLPOST--LFRADHPFLFVIR--KDDIILFSGKXSCP 380
 356 ---PTSTNEEFCHNHPFLFIRQNTKNSILFYGRFSSP 390

Db 61 DTN-----EDHSRFQSLNADINKGAPYILKLANRYGKTYNEFLADFLASTOK 110

Oy 121 LDYAKVERVDFTNHLDETRRNINKWVENETHGKIKVYIGGGISSAVLVNAYVYKRG 180

Db 111 MYGAEIASVDFQAPPEAKKEINEMVKGTEBKIPELLYGMDNMKTLVNAIYKGN 170

Oy 181 MOSAFKSETINCHFKSPKSGKAVAMHOERKFNLSVIDPSKKILELYNG-GIMMYV 239

Db 171 MOEFKMEATRDAPFRLNKOTFTVKMMYQKKKFPYVIEDLKCRLVELPFOGKELSMI 230

Oy 240 LLEPN-----DLSEINKTLFONLMEWTPRRMTSKYVEFFPOFKIEKNYKQYLRA 293

Db 231 LRPDDLEDESTGLEKIEKQLTLEKLRMTKPEMLYLAENVVHLPRFKLEESYDLTSHLAR 290

Oy 294 LGLKDIDESKADLSGIASGRLYISRMHKSYTEVEEGTEATGATGNSINVEKQLPOST 353

Db 291 LGVODLFNRKADLSGSGARDLFVSKTIHKSFDVLEEGTEAATAAGTATMLAMPPEE 350

Oy 354 LFRADHPFLFVIRKDI--ILFSGKVSQP 380

Db 351 NFNADHPFIFIRHNPSANILFLGRFSSP 379

RESULT 4

elastase inhibitor - human

C:Species: Homo sapiens (man)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Sep-1997

C:Accession: S27383; S65750

R:Remold-O'Donnell, E.; Chin, J.; Alberts, M.

Proc. Natl. Acad. Sci. U.S.A. 89, 5635-5639, 1992

A:Title: Sequence and molecular characterization of human monocyte/neutrophil elastase

A:Reference number: S27383; MUID:92302296; PMID:1376927

A:Accession: S27383

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-379 <REM>

A:Cross-references: EMBL:M03056

R:Packard, B.2.; Lee, S.S.; Remold-O'Donnell, E.; Komoriya, A.

Biochim. Biophys. Acta 1269, 41-50, 1995

A:Title: A serpin from human tumor cells with direct lymphoid immunomodulatory activity:

A:Reference number: S65750; MUID:96049524; PMID:7578269

A:Accession: S65750

A:Status: preliminary

A:Molecule type: protein

A:Residues: 57-65;97-110;111-129;204-213;216-244;255-271,'X',273-274;291-301 <PAC>

C:Genetics:

A:Gene: GDB:ELANH2; EI; P12

A:Cross-references: GDB:132914; OMIM:130135

A:Map position: 6p25-6p24.3

C:Superfamily: antithrombin III

C:Keywords: serine proteinase inhibitor

Query Match 36.9%; Score 722.5; DB 2; Length 379;

Best Local Similarity 38.6%; Pred. No. 1.5e-43;

Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

Oy 1 MASLAANAEPFCFLFREMDDNGNGVFPSSLSLFAALALVRLGADDSLSQIDKLHV 60

Db 1 MEOLSSANTRFADLFLALSENPNAGNIFISPFISAMAMVFLGTGNTAAOQSKTFHF 60

Oy 61 NTASGYCNSSNSGCSQSLKRVFSDINASHKVDLSIVNGLFAEKYGGFKKDIIECAEK 120

Db 61 NTV-----EVSRSQSLNADINKGASITLKLANRLYGEKTYNFLPEFLVSTOK 110

Oy 121 LDYAKVERVDFTNHLDETRRNINKWVENETHGKIKVYIGGGISSAVLVNAYVYKRG 180

Db 111 TYGADLASVDFOHASEDAKRTINQWVGQTEGKIPELLASGMVDMNMTKIVLVNAIYKGN 170

Oy 181 MOSAFKSETINCHFKSPKSGKAVAMHOERKFNLSVIDPSKKILELYNG-GIMMYV 239

Db 171 WKDFKMEATRDAPFRLNKDKRTKVMYOKKFKFAYGIEDLKCRVLELPFOGKELSMI 230

Oy 240 LLEPN-----DLSEINKTLFONLMEWTPRRMTSKYVEFFPOFKIEKNYKQYLRA 293

Db 231 LRPDDLEDESTGLEKIEKQLTLEKLRMTKPEMLYLAENVVHLPRFKLEESYDLTSHLAR 290

Oy 294 LGLKDIDESKADLSGIASGRLYISRMHKSYTEVEEGTEATGATGNSINVEKQLPOST 353

Db 291 LGVODLFNRKADLSGSGARDIFISKIVHKSFEVNEEGTEAATAAGTATFCMLPPEE 350

Oy 354 LFRADHPFLFVIRKDI--ILFSGKVSQP 380

Db 351 NFNADHPFIFIRHNPSANILFLGRFSSP 379

RESULT 5

headpin serine proteinase inhibitor - human

C:Species: Homo sapiens (man)

C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 26-May-2000

C:Accession: JC7118

R:Spring, P.; Nakashima, T.; Frederick, M.; Henderson, Y.; Clayman, G.

Biochem. Biophys. Res. Commun. 264, 299-304, 1999

A:Title: Identification and cDNA cloning of headpin, a novel differentially expressed

A:Reference number: JC7118; MUID:99458651; PMID:10527881

A:Accession: JC7118

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-391 <SPR>

A:Cross-references: GB:AF169949; NID:95911368; PIDN:AAD55765.1; PID:95911369

C:Genetics:

A:Map position: 18q21.3-18q22

C:Superfamily: antithrombin III

C:Keywords: serine proteinase inhibitor

F:355/Inhibitory site: Val (unidentified proteinase) #status predicted

Query Match 36.5%; Score 715.5; DB 2; Length 391;

Best Local Similarity 38.0%; Pred. No. 4.8e-43;

Matches 152; Conservative 78; Mismatches 141; Indels 29; Gaps 7;

Oy 1 MASLAANAEPFCFLFREMDDNGNGVFPSSLSLFAALALVRLGADDSLSQIDKLHV 60

Db 1 MDSIGAAYSTRUGFLFKEL-KRTNDGNIFESPVGILTAIGVLLGTGATASOLEEVFH- 58

Oy 61 NTASGYCNSSNSGCSQSLKRVFSDINASHK-----DYDSIVNGLPAEK 106

Db 59 -----SEKFKSSRIAAEKREYIENTEAVHQOQKLTESKLTNDYELNTRKLGK 112

Oy 107 VYGFHKDYIRCOAEKLYDAKVERVDFTNHLDETRRNINKWVENETHGKIKVYIGGGISS 166

Db 113 TYFLQKLYLDVEKRYHASLEPVDVFNAADESRKINSWESKTEKIKDLFPGSSISS 172

Oy 167 AVMYLVNAYVFKGKWSAFKSETINCHFKSPKSGKAVAMHOERKFNLSVIDPSMKI 226

Db 173 TKLVNMYVYFGQMDREFKKEKFMNKNSTKSVOYMTOSHSFSFTELDQAKI 232

Oy 227 LELKY-NGGINMYVLLRPNDLSEITE---NKLTFLONLMEWTPRRMTSKYVEFFPOFKIE 282

Db 233 LGIYKKNNDLSMFLL-NDIDGLEKTIIDKISPKLVEMTSPGMEERKYNLHPREVE 291

Oy 283 KNYEMQYLRALGKIDFESKADLSGIASGRLYISRMHKSYTEVEEGTEATGATGNSINVEK 342

Db 292 DSYDLAVNLAMGQDAFSEHKADYSGMSSGGLYAKFLHSSFVAVTEGTEAATAAGT 351

Oy 343 NIVEKQLPOSTLFRADHPFLFVR--KODIILFSGKVSQP 380

Db 352 GFTVTSAPGHENVCHNHPFLFIFIRHNPSNITLFGRESSP 391

RESULT 6

DYCH ovalbumin-related y protein - chicken

C:Species: Gallus gallus (chicken)

C>Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 18-Jun-1999

C:Accession: A01244

A:Cross-references: GB:M18082; NID:g1895562; PIDN:AAA60006.1; PID:g189563
 R:Ye, R.D.; Mun, T.C.; Sadler, J.E.
 J. Biol. Chem. 262, 3718-3725, 1987
 A:Title: cDNA cloning and expression in Escherichia coli of a plasminogen activator inh
 A:Reference number: A26553; MUID:87137674; PMID:3029122
 A:Accession: A26553
 A:Molecule type: mRNA
 A:Residues: 1-119, 'D', 121-403, 'K', 405-412, 'C', 414-415 <YE2>
 A:Cross-references: GB:J02665; NID:g1895544; PIDN:AAA36413.1; PID:g189545
 A:Experimental source: Placenta
 R:Jensen, P.H.; Schueler, E.; Woodrow, G.; Richardson, M.; Goss, N.; Hojrup, P.; Petersen
 J. Biol. Chem. 269, 15394-15398, 1994
 A:Title: A unique interhelical insertion in plasminogen activator inhibitor-2 contains t
 A:Reference number: A53815; MUID:94253109; PMID:7910824
 A:Accession: A53815
 A:Molecule type: protein
 A:Residues: 83-87 <JEN>
 R:Kruthof, E.K.O.; Vassalli, J.D.; Schleuning, W.D.; Mattaliano, R.J.; Bachmann, F.
 J. Biol. Chem. 261, 11207-11213, 1986
 A:Title: Purification and characterization of a plasminogen activator inhibitor from the
 A:Reference number: A25021; MUID:86278222; PMID:3050045
 A:Accession: A25021
 A:Molecule type: protein
 A:Residues: 347-376 <KR>
 R:Kiso, U.; Kaudevitz, H.; Henschen, A.; Astedt, B.; Kruthof, E.K.O.; Bachmann, F.
 FEBS Lett. 230, 51-56, 1988
 A:Title: Determination of intermediates, products and cleavage site in the reaction betw
 A:Reference number: S02435; MUID:88167197; PMID:3280346
 A:Accession: S02435
 A:Molecule type: protein
 A:Residues: 377-415 <KIS>
 R:Kruthof, E.K.; Cousin, E.
 Biochem. Biophys. Res. Commun. 156, 383-388, 1988
 A:Title: Plasminogen activator inhibitor 2. Isolation and characterization of the promot
 A:Reference number: I52229; MUID:89025873; PMID:2845977
 A:Accession: I52229
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-56 <RE2>
 A:Cross-references: GB:M23092; NID:g189559; PIDN:AAA60005.1; PID:g189561
 C:Genetics:
 A:Gene: GDB:PA12; PLANH2
 A:Cross-references: GDB:120298; OMIM:173390
 A:Map position: 18q21.2-18q22
 A:Insertions: 56/3; 96/3; 139/3; 179/1; 226/3; 281/3
 C:Superfamily: antithrombin III
 C:Keywords: acetylated amino end; glycoprotein; serine proteinase inhibitor
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-415/Product: plasminogen activator inhibitor 2 #status predicted <MAT>
 F:23/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:75,115,339/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:83,84,86/Cross-link: Isopeptide (Gln) (interchain to Lys N6-amino of unidentified prot
 F:380-381/Cleavage site: Arg-Thr (plasminogen activator) #status experimental
 F:380/inhibitory site: Arg (plasminogen activator) #status experimental

Query Match 33.9%; Score 664.5; DB 2; Length 415;
 Best Local Similarity 34.7%; Pred. No. 2,1e-39;
 Matches 144; Conservative 83; Mismatches 153; Indels 35; Gaps 5;

1 MASLAANAEEFCNLFREMDNONGCNVFFSSLSLFAALALVLRGADDSLSQDKLLHV 60
 1 MEDLCVANTLFLALNLFHLAKASPTONLFLSPMSISSTMAVYWGSGSTEDONAKVLOF 60
 61 N-----TASGY-----GN-----SSNSQSGIOSQLKRVFSDINASHKD 93
 61 NEVGANAVTPTMPTENPFCGFMOOIKGSPDALIQAOADKIHSSFRSLSSAINASTGN 120
 94 YDLSIVNGLFAEKYGGHXYGIBCAEKLYDAKVERVDFTHNLEPTRRINIKKVVNENHGK 153
 121 YLLESVAKLFGKESASRREETIRLCQYYSSEPAVDFLECAEARKKINSWVATQKCK 180
 154 IKNVIGGGISSAVVNLVNAVYFKGMOAFTKSETINCHFKSPKSCGKAVAMHDERK 213

181 IPNLLPEGSVDGRTAVLVNAVYFKGKMTKPFKEKLLNGLYPFRVNSAORTPVOMYLRER 240
 214 FNLSVIEDPSPKMTIELRVNGINMYVLPE-----NDLSEIEKTLFPOUMLEKTNRRM 267
 241 LNTGVIDLKAQILIELPAGDVSMFLLPDLADVSTGLLESELYDILKNTKTSXDKM 300
 268 TSKVVEFPFOFKIEKNYENKQYLRALGLKDIPEDSKADSLASGGRLYISRMHNSYI 327
 301 AEDEVEYIYQFKIEHTELRSLRSKMGMEDAFKKGANSGMSERNDLFLSEVFRQAVY 360
 328 EYTEEGTEAPATGSGNIVERKQLPOSTLFRADHPELVY--RKDDIILFSGKVSQP 380
 361 DVNEEGTEAAGTGVMTGRTGHGQPGVADHPEFLIMKTRICILFGRFSSP 415

RESULT 9
 A59273
 N:Alternate names: cytoplasmic antiprotease 2 (CAP-2)
 C:Species: Homo sapiens (man)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: A59273
 R:Speicher, C.A.; Morgenstern, K.A.; Mathewes, S.; Dahlen, J.R.; Schrader, S.K.; Fost
 J. Biol. Chem. 270, 29854-29861, 1995
 A:Title: Molecular cloning, expression, and partial characterization of two novel mem
 A:Reference number: A59273; MUID:96102039; PMID:8550382
 A:Accession: A59273
 A:Status: preliminary; not compared with conceptual translation
 A:Cross-references: GB:I40377; NID:g1160926; PIDN:AAC41939.1; PID:g1160927
 A:Residues: 1-374 <SPR>
 A:Molecule type: mRNA
 A:Cross-references: GB:I40377; NID:g1160926; PIDN:AAC41939.1; PID:g1160927
 A:Experimental source: tissue type placenta; note (vector lambda gtl1); gene CAP2
 C:Genetics:
 A:Gene: GDB:PI8; CAP2
 A:Cross-references: GDB:599392; OMIM:601697
 A:Map position: 18q21.3-18q21.3
 C:Superfamily: antithrombin III
 C:Keywords: proteinase inhibitor
 F:339/inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match 33.8%; Score 662; DB 2; Length 374;
 Best Local Similarity 37.4%; Pred. No. 2,7e-39;
 Matches 145; Conservative 79; Mismatches 142; Indels 22; Gaps 8;

1 MASLAANAEEFCNLFREMDNONGCNVFFSSLSLFAALALVLRGADDSLSQDKLLHV 60
 1 MDLCVANTLFLALNLFHLAKASPTONLFLSPMSISSTMAVYWGSGSTEDONAKVLOF 60
 61 NTASGYGNSSNSQSGIOSQLKRVFSDINASHKDYDLSIVNGLFAEKYGGHXYGIBCAEK 120
 61 Y-----KDGDIRHGFOSL-----SEVNRGTQVLLTANRLREGKTCDFLPQKCYCK 110
 121 LYDAKVERVDFTHNLEPTRRINIKKVVNENHGKIKNYIGSGGSISSAVVNLVNAVYFKG 180
 111 FYQAELEBELSFAEDTECRKHNDVAKTEGKISEVLDAGTVDPLKLVNAVYFKG 170
 181 WQSAFTKSETINTCHFKSPKSCGKAVAMHDERKNSLVEDPSKILFLRY--NGITMYV 239
 171 WNEQFDRKRYTGMIFKTNE--EKTVMOMFEAKKMOYADEVHTQVLELPEVEEELSMVI 229
 240 LLEPEN--DLSEIEKTLFPOUMLEWTNPRRTSKVVEFPFOFKTEKVEKQYLRALGLK 297
 230 LLEPDNTDLAVKALTYEKFKANTNSEKILTSKVQVFLRLKLEESTDLEPLRLRIGMT 289
 298 DIPEDSKADSLASGGRLYISRMHNSYIETVEGTEAATG--SNIVERKQLPOSTL 354
 290 DAFDEAKADPSGMSKTEKVNPLSKVAHRCFEVNEEGTEAATAAVVNSNSCSMRER--- 346
 355 FRADHPELVYR--KDDIILFSGKVSQP 380
 347 FCADHPELVYRIRRKTKCILFCGRFSSP 374

Query Match	33.6%	Score 658.5	DB 2	Length 383
Best Local Similarity	35.0%	Pred. No. 4.9e-32		
Matches 137	Conservative 96	Mismatches 139	Indels 19	Gaps 7
QY	1	MASLAANAABECFLNLFREDDNOGNGVNFSSLSLFAALALATVRLGADDSLSQIDKLHLV	60	
Db	1	MSGISAGSMEECFDFVFKELKVHANNDMLSPFALSTLWVFLGAKOSTRTQIKVYHF	60	
QY	61	NTASGCGSSNSQSG----LOSQKRVSDINASKDQDLSVNGLSVLAFAEKYVGFKKDYE	116	
Db	61	DKLPQFGSIEAOGCTSVNVHSSLSLDLINOITKQDASFSLSASLYAOETTVVPEYLQ	120	
QY	117	CAEKLVDKAYVERVDYFNHLEDTRNRINKVKNVENETHGKIKVNGIGEGSISSAVMYLVNAVY	176	
Db	121	CVKELVRCGLSESVNQTADQADARGILNMVSESQTGIIIRNLIPQSVBSQTFMYLVNMAIA	180	
QY	177	FKGKMQSAFTKSETINCHFKSPKCSGCAVAMHMRKFNLSVIEDPSMKILELRY-NGCI	235	
Db	181	FKGLMEKFRKAEEDPOTIPFRVTEDESKPVOMVOIGTSFKVASMASEKKKILELPRASGT	240	
QY	236	NMVYLPR--NDLSIEIKKLFQONLMEKNTNPRBMKSVKEVPEPFQKIEKNVEMQOYRA	293	
Db	241	SMVLVLLPDDVSGLEOLEEIIISFEKILETWTSSIMERVKVYILPRKMKBEKYNLSILMA	300	
QY	294	LGALNIDESKADSLGSIAGRLVYSRMKHYSYLEVEEGTE--ATAATGSNIVEKOLPO	351	
Db	301	MGINDLF--SSANLSGISISVSLKISQAVVHAHAHETNAGRDVGSAAEGVDATEE----	355	
QY	352	STLRADHPLFLVIR--KDDIILSPGAKSCP	380	
Db	356	---FRADHPLFCVAKHETNALILPGRCVSP	383	

RESULT 12
I39184
bomaplin - human
N:Alternate names: proteinase inhibitor 10 (PI10)
C:Species: Homo sapiens (man)
C:Date: 01-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 16-Jul-1999
C:Accession: I39184
R:Riewald, M., Schlieef, R.R.
J. Biol. Chem. 270, 26754-26757, 1995
A:Title: Molecular cloning of bomaplin (proteinase inhibitor 10), a novel human serpin that
A:Reference number: I39184; MUID:96070759; PMID:7592909
A:Accession: I39184
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-397 <RES>
A:Cross-references: EMBL:U35459; NID:g1065408; PIDN:AAC50282.1; PID:g1065409
C:Genetics:
A:Gene: GDB:PI10
A:Cross-references: GDB:636283
A:Map position: 14q32.1-14q37.1
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor

	Query Match Similarity	33.6%	Score 658.5	DB 2:	Length 397;
	Best Local Similarity	34.9%	Pred. No.51e-39;		
	Matches 139;	Conservative 89;	Mismatches 151;	Indels 19;	Gaps 5
Oy	1 MASIAAANAEECFNLFRMEDDNOGNGVNFSSLSFLALALVRLOAOODSLSDIKLHV	60			
Dd	1 MDLSLATISNQALETSKKLAESAGCKNIFFSSWSTLTETIYLLCAKGTTAAQMAQVLOF	60			
Oy	61 NTAASGYC-----NSSNSQSGLSQOLAKVFESDINASHKDYDLSIVNGLFAEYK	107			
Dd	61 NRDOGVKCDPSESEKKRKMEPLNSNEE-IHSPFTLLISEILKPNDYLLKTANAIYGEXT	119			
Oy	108 YGFHKDYTECAEKLYDAKVERVDFTFNHLIEDPRRRINKKVENVETHRKKIKNVIGEGGISSA	167			
Dd	120 YAFHNKKYTEDKMTYFGAEPQVPNPFEASDAQIRKIDINSVVERTKEKIONLLDDDSVDSTT	179			

[illegible]

RESULT 13
A57488
proteinase inhibitor Sp13 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 16-Jul-1999
C:Accession: A57488
R:Sun, J.; Rose, J.B.; Bird, P.
J. Biol. Chem. 270, 16089-16096, 1995
A:Title: Gene structure, chromosomal localization, and expression of the murine homol
A:Reference number: A57488; MUID:95332310; PMID:7608171
A:Accession: A57488
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <SUN>
A:Cross-references: GB:U25844; NID:g818902; PIDN:AAA79684.1; PID:g818903
A:Genetics:
A:Gene: Sp13
A:Map position: 13
C:Superfamily: antithrombin III
C:Keywords: serine proteinase inhibitor
C:343/inhibitory site: Arg (unidentified proteinase) #status predicted

Query Match	33.1%	Score 649;	DB 2;	Length 378;
Best Local Similarity	37.9%	Pred. No. 2,2e-38;		
Matches	146;	Conservative	77;	Mismatches 150; Indels 12; Gaps 7
QY	1	MASLAANAEEFCNLFREMDNQCNGNVFSSLSFLAALALVRLGAQDDSLQIDKLHV	60	
DB	1	MDPLQEAAGTALNLTLLKILGEDSSK - NYFLSPMSISSIALAMFAMKGTASOMAOAL	59	
QY	61	NTAAGYGSSNSOSGLOSOLKRVPSDINAHKDYDLSVNGLFAPKAYGVGHNDYTECAK	120	
DB	60	DKCSGNGGGDVHQ - GFOSLL - - - TEVAKTGTQYLLRTANRLEFGKTCDDLASFSDSCLK	114	
QY	121	LYDAKVERVEFTNHLIEDTRRNINKVVENETHGKIKNVIGEGGISSAAVVLNAVAYFGK	180	
DB	115	FYFAELIELDPOGATEESRQHINFWVAKKTDEKIEVLSPTGVNSDTSLVNALATYFGN	174	
QY	181	WOSAFYKSETINCHFSPKCSGKAKAVAMNHQEFKFLSYIEDPSMKILBARY - NGCINMY	239	
DB	175	WEKOFNKENTEMPEPKSKNEEKPYOMAFKSTFMTYIGELFTSLILLPYVSSBLMII	234	
QY	240	LLPEN - DLSEIENKLTGFONLMEHTNPPRMYSKYVEVFPOPKIEKNEMKOYALBGLK	297	
DB	235	MLPDEHVLISVEKEVETIEKRIETWRDLKMDDEEVEVLPRKKLEENTNMNDALTKLGMT	294	
QY	298	DIFDESKADLSGLASGRRLYISRMHKSXYIEVTEEGTEATATGSGNIVEKOLPOSTLRLA	357	
DB	295	DAFG - GRADFGSMSSKQGLFLSKVYHNKAFVEVNEEGTEAALAAATAGMTVMCRIFPRPCA	353	
QY	358	DHPLFLVLR - - KDQILFLSGKVSCP	380	
DB	354	DHPLFLFTYHNWKTNGILFCGRSSS	378	

RESULT 14

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 36.341 Seconds

(Without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-34
Perfect score: 1959
Sequence: 1 MASLAANAEFCNLFREMD.....FLFVIRKDDILFSGKVSQP 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	1959	100.0	380	1 SPB7_HUMAN	O75635 homo sapien
2	819	41.8	392	1 SBI1_HUMAN	O96p15 homo sapien
3	763.5	39.0	405	1 SBI2_HUMAN	O96p63 homo sapien
4	752	38.4	390	1 SCC1_HUMAN	P29508 homo sapien
5	748	38.2	390	1 SCC2_HUMAN	P48594 homo sapien
6	739.5	37.7	379	1 ILEU_HORSE	P05619 equus caball
7	722.5	36.9	379	1 ILEU_HUMAN	P30740 homo sapien
8	714.5	36.5	391	1 SBI3_HUMAN	O9u1v8 homo sapien
9	711	36.3	388	1 OVAY_CHICK	P01014 gallus gall
10	702	35.8	378	1 PTIG_BOVIN	O02739 bos taurus
11	696	35.5	378	1 ILEU_PIG	P80239 sus scrofa
12	664.5	33.9	415	1 PAI2_HUMAN	P05120 homo sapien
13	662	33.8	374	1 SPB8_HUMAN	P50452 homo sapien
14	658.5	33.6	397	1 SBI0_HUMAN	P48595 homo sapien
15	654	33.4	385	1 OVAL_CHICK	P01012 gallus gall
16	653.5	33.4	382	1 OVAL_COTJA	P19104 coturnix co
17	649	33.1	378	1 PTIG_MOUSE	O60854 mus musculu
18	644	32.9	376	1 PTIG_HUMAN	P35237 homo sapien
19	630.5	32.2	415	1 PAI2_MOUSE	P12338 mus musculu
20	623	31.8	376	1 SPB9_HUMAN	P50453 homo sapien
21	622	31.8	416	1 PAI2_RAT	P29524 rattus norv
22	605.5	30.9	375	1 MASP_RAT	P70124 rattus norv
23	600.5	30.7	375	1 MASP_MOUSE	P32261 mus musculu
24	592	30.2	465	1 ANT3_MOUSE	P32261 mus musculu
25	590.5	30.1	375	1 MASP_MOUSE	P32261 mus musculu
26	567	28.9	433	1 ANT3_BOVIN	P36932 homo sapien
27	566	28.9	464	1 ANT3_HUMAN	P41361 bos taurus
28	557	28.4	454	1 ANT3_SHEEP	P12008 homo sapien
29	503.5	25.7	410	1 NEUS_RAT	P32262 ovis aries
30	498.5	25.4	410	1 NEUS_HUMAN	O91142 rattus norv
31	492	25.1	410	1 SP12_MOUSE	O99574 homo sapien
32	484.5	24.7	410	1 NEUS_MOUSE	O91k88 mus musculu
33	478.5	24.4	417	1 KBP_MOUSE	P29621 mus musculu

34	476.5	24.3	410	1 NEUS_CHICK	O90935 gallus gall
35	476.5	24.3	423	1 AACT_HUMAN	P01011 homo sapien
36	467	23.8	405	1 SP12_HUMAN	O75830 homo sapien
37	467	23.8	410	1 ALAT_DIDMA	O03044 didelphis m
38	462.5	23.6	418	1 COTR_MOUSE	P07739 mus musculu
39	460	23.5	418	1 S124_APOXY	O60336 apodemus sy
40	457.5	23.4	413	1 ALMS_TAMSI	O54758 tamias sibi
41	455	23.2	418	1 CP16_RAT	P09006 rattus norv
42	451	23.0	405	1 IPSP_MOUSE	P70458 mus musculu
43	446.5	22.8	405	1 ALAS_CAVPO	P22325 cavia porce
44	446.5	22.8	406	1 IPSP_HUMAN	P05154 homo sapien
45	444.5	22.7	413	1 ALMM_TAMSI	O54757 tamias sibi

ALIGNMENTS

RESULT 1
ID SPB7_HUMAN STANDARD; PRT; 380 AA.
AC O75635;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Megsin (TPP5) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97326116; PubMed=9182567;
RA Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iwasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity".
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98376492; PubMed=9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.,
RT "A mesangium-predominant gene, megsln, is a new serpin upregulated in
RT J. Clin. Invest. 102:828-836(1998).
CC -!- FUNCTION: Might influence the maturation of megakaryocytes via
CC proteases. Might influence the maturation of Lys-specific
CC its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D88575; BAA31232.1; -
CC EMBL: AF027866; AAC64506.1; -
CC HSSP: P05619; IHLF.
CC Genew: HGNC:13902; SERPINB7.
CC MTM: 603357; -
CC InterPro: IPR000215; Serpin.
CC Pfam: PF00079; serpin; 1.
CC SMART: SM00093; SERPIN; 1.
CC PROSITE: PS00284; SERPIN; 1.

KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 347 348 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C3CFF605 CRC64;

Query Match 100.0%; Score 1959; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1,2e-129;
 Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MASLAANAEFCENLREMDNDNGNGNVEFSSLSLFAALALVRLGAODSLSDIKLLHY 60
DB 1 MASLAANAEFCENLREMDNDNGNGNVEFSSLSLFAALALVRLGAODSLSDIKLLHY 60
OY 61 NTASGCGNSNSGSLQSLKRVESDINASHKDYDLSYNGLFAEKVYGFHKDYTECAEK 120
DB 61 NTASGCGNSNSGSLQSLKRVESDINASHKDYDLSYNGLFAEKVYGFHKDYTECAEK 120
OY 121 LVDAKERVDFNTHLEDTRRNINKWVENETHGKIKRVIGEGGSSSAVAVLVNAVYFKRK 180
DB 121 LVDAKERVDFNTHLEDTRRNINKWVENETHGKIKRVIGEGGSSSAVAVLVNAVYFKRK 180
OY 181 WQSAPFTSEITNCHFKSPKCSGKAVAMHQRKFNLSYEDPSMKLLEIRYNGIMAYVL 240
DB 181 WQSAPFTSEITNCHFKSPKCSGKAVAMHQRKFNLSYEDPSMKLLEIRYNGIMAYVL 240
OY 241 LPENDLSEIENKLTFRONLMEWTNPRRMYSKYVEFPPOFKIEKNEMKOYLKALGKIDF 300
DB 241 LPENDLSEIENKLTFRONLMEWTNPRRMYSKYVEFPPOFKIEKNEMKOYLKALGKIDF 300
OY 301 DESKADLSGIASGGRLYISRMHKSXYIEVTEGETATGSIYVKOLPOSTLFRADHP 360
DB 301 DESKADLSGIASGGRLYISRMHKSXYIEVTEGETATGSIYVKOLPOSTLFRADHP 360
OY 361 FLFVIRKDDIILFSGKVSCEP 380
DB 361 FLFVIRKDDIILFSGKVSCEP 380

```

RESULT 2

```

SB11_HUMAN STANDARD; PRT; 392 AA.
ID SB11_HUMAN 096P14; 096P13;
AC 096P15; 096P14; 096P13;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serpin B11.
GN SERPINB11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS THR-148 AND THR-181.
RA Silverman G.A., Cataltepe S.;
RT "SERPINB11 is a novel ov-serpin with at least 3 alleles."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF419953; AAL16056.1; -
DR EMBL: AF419954; AAL16057.1; -
DR EMBL: AF419955; AAL16058.1; -
DR GenBank: HGNC:14221; SERPINB11.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; Serpin.1.
DR PROSITE: PS00284; SERPIN.1.

```

KW Serpin; Serine protease inhibitor; Polymorphism.
 FT ACT_SITE 357 358 REACTIVE BOND (BY SIMILARITY).
 FT VARIANT 148 148 M -> T (IN ALLELE B).
 FT VARIANT 181 181 /FTID-VAR_012472
 FT VARIANT 181 181 I -> T (IN ALLELE B AND ALLELE C).
 FT /FTID-VAR_012473.
 SQ SEQUENCE 392 AA; 44098 MW; 906F6DD412BCD756 CRC64;

Query Match 41.8%; Score 819; DB 1; Length 392;
 Best Local Similarity 43.1%; Pred. No. 4.1e-50;
 Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

```

OY 1 MASLAANAEFCENLREMDNDNGNGNVEFSSLSLFAALALVRLGAODSLSDIKLLHY 60
DB 1 MASLAANAEFCENLREMDNDNGNGNVEFSSLSLFAALALVRLGAODSLSDIKLLHY 60
OY 61 NTASGCGNSNS-SQSG-LQSLKRVESDINASHKDYDLSYNGLFAEKVYGFHKD 113
DB 61 NTASGCGNSNS-SQSG-LQSLKRVESDINASHKDYDLSYNGLFAEKVYGFHKD 113
OY 114 YTECAEKLVDAKERVDFNTHLEDTRRNINKWVENETHGKIKRVIGEGGSSSAVAVLVN 173
DB 114 YTECAEKLVDAKERVDFNTHLEDTRRNINKWVENETHGKIKRVIGEGGSSSAVAVLVN 173
OY 121 YLSCSEKMYQALQIVDFEOSTEETRKMNAVENKTNCKVANLFGKSTIDPSSVAVLVN 180
DB 121 YLSCSEKMYQALQIVDFEOSTEETRKMNAVENKTNCKVANLFGKSTIDPSSVAVLVN 180
OY 174 AVYFGKQWQSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSYEDPSMKLLEIRY-N 232
DB 174 AVYFGKQWQSAFTKSETTINCHFKSPKCSGKAVAMHQRKFNLSYEDPSMKLLEIRY-N 232
OY 233 GGNNVYVLLEN--DLSEIENKLTFRONLMEWTNPRRMYSKYVEFPPOFKIEKNEMKOY 290
DB 233 GGNNVYVLLEN--DLSEIENKLTFRONLMEWTNPRRMYSKYVEFPPOFKIEKNEMKOY 290
OY 291 LKALGLKIDFDESKADLSGIASGGRLYISRMHKSXYIEVTEGETATGSIYVKOLP 350
DB 291 LKALGLKIDFDESKADLSGIASGGRLYISRMHKSXYIEVTEGETATGSIYVKOLP 350
OY 301 LKPLGVTDLFNQYKADLSGMSPTKGLYLSKALHKTYLDVSEGTBAATGDSIAVKSLP 360
DB 301 LKPLGVTDLFNQYKADLSGMSPTKGLYLSKALHKTYLDVSEGTBAATGDSIAVKSLP 360
OY 361 OSTLFRADHPFLFVIR--KDDIILFSGKVSCEP 380
DB 361 OSTLFRADHPFLFVIR--KDDIILFSGKVSCEP 380

```

RESULT 3

```

SB12_HUMAN STANDARD; PRT; 405 AA.
ID SB12_HUMAN 096P63;
AC 096P63;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serpin B12.
GN SERPINB12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA MEDLINE-21623618; PubMed-11604408;
RA Askew Y.S., Pak S.C., Luke C.J., Askew D.J., Cataltepe S., Mills D.R.,
RA Kato H., Lebowitz J., Dewar K., Birren B., Silverman G.A.;
RT "SERPINB12 is a novel member of the human ov-serpin family that is
RT widely expressed and inhibits trypsin-like serine proteinases."
RL J. Biol. Chem. 276:49320-49330(2001).
CC -1- FUNCTION: Inhibits trypsin and plasmin, but not thrombin,
CC coagulation factor Xa, or urokinase-type plasminogen activator.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in many tissues, including brain,
CC bone marrow, lymph node, heart, lung, liver, pancreas, testis,
CC ovary, and intestine.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF11191; AL05571.1; -
DR Genew: HGNC:14220; SERPINB12.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR PROSITE: PS00284; SERPIN.1.
KM Serpin: Serine protease inhibitor.
FT ACT_SITE 370 371 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 405 AA; 46276 MW; FFE12D4C9B7F3DPA CRC64;

Query Match 39.0%; Score 763.5; DB 1; Length 405;
Best Local Similarity 39.3%; Pred. No. 3.2e-46;
Matches 155; Conservative 87; Mismatches 114; Indels 25; Gaps 5;

QY 1 MASLAANAEEFCNLFREMDNOGNVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
DB 1 MDSLVTANTKFCFDLFOEIGKDRHKNIFFSPILSLAALGMVRLGARSDSAHOIDEVLHF 60
QY 61 NTAS-----GYNSSNSQGLQG-OLKRVPSDINASHKDYDLTVNGLEFA 104
DB 61 NEFSQNSKEPDPCLKSNKOKAGSLNNSGLVSCYFQGLSKLDRIKTDYTLSTANRLYG 120
QY 105 EKVYGFHKDYIECAEKLYDAKVERVDFTNLEDRRNINRWENETHGKIKNYEGEGIS 164
DB 121 EGEFPICQETVDGYOYHTTIESVDQKNPKERSKQENWECOSQKIKELFSKALIN 180
QY 165 SSAVMVLNVNVYFKGQKSAFTSKSETINCHFKSPKSGKAVAMHQRKNLSVIEDPSM 224
DB 181 AEVLVLNVNVYFKAKETFEYFHNENTADAPCLANENKSKVMKTKGLRIGIEVKA 240
QY 225 KILELRY-NGGINMYVLLPND-----LSEIENKLTFOUMLEMTNPRRTSKYVEVEFP 277
DB 241 QILEMRYTKGLKSLFVLLPSSHDKNLKGLELEKRIYERKVMVAASSSENVSVLFP 300
QY 278 OFRIEKRYEMKQYLRALGLDIFDESKADLSGIASGRLYISRMHKSYLEVTEGTEAT 337
DB 301 RFLLEDSTDLNSLQDDGIDTIDETRADLGLSPSPNLTKLTHKTFVEVDENGTOAA 360
QY 338 AATGSNIVERKOLPOSTLEFRADHPFLFYIR--KDDILFSGKVSQP 380
DB 361 AATGAVAVSERLSRWSWFEFNNHPFLFIRHNKTQTLIFGYRVCSF 405

RESULT 4
SCCL_HUMAN
ID SCCL_HUMAN STANDARD; PRT; 390 AA.
AC P29508: 096321;
DT 01-APR-1993 (Rel. 25, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Squamous cell carcinoma antigen 1 (SCCA-1) (Protein T4-A).
GN SERPINB3 OR SCCA1 OR SCCA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=92068241; Pubmed=1958219.
RA Suminami Y., Kishi F., Sekiguchi K., Kato H.;
RT "Squamous cell carcinoma antigen is a new member of the serine
RT protease inhibitors."
RL Biochem. Biophys. Res. Commun. 181:51-58(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95241462; Pubmed=7724531.
RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,
RA Treter S.D., Hui S.M., Silverman G.A.;
RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem

RT duplication of the human squamous cell carcinoma antigen gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
CC IMMUNE RESPONSE AGAINST TUMOR CELLS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.
CC -1- DEVELOPMENTAL STAGE: ITS EXPRESSION IS CLOSELY RELATED TO CELLULAR
CC DIFFERENTIATION IN BOTH NORMAL AND MALIGNANT SQUAMOUS CELLS.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL: S66896; AAB20405.1; -
DR EMBL: U19556; AAA97552.1; -
DR EMBL: U19568; AAA86317.1; -
DR EMBL: U19559; AAA86317.1; JOINED.
DR EMBL: U19560; AAA86317.1; JOINED.
DR EMBL: U19565; AAA86317.1; JOINED.
DR EMBL: U19567; AAA86317.1; JOINED.
DR EMBL: U19562; AAA86316.1; JOINED.
DR EMBL: U19559; AAA86316.1; JOINED.
DR EMBL: U19560; AAA86316.1; JOINED.
DR EMBL: BC005224; AAH05224.1; -
DR PIR: J10966; J10966.
DR HSSP: P01008; 1ATH.
DR Genew: HGNC:10569; SERPINB3.
DR MIM: 600517; -
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART: SM00093; SERPIN.1.
DR PROSITE: PS00284; SERPIN.1.
KM Serpin: Serine protease inhibitor.
FT ACT_SITE 354 355 REACTIVE BOND.
FT CONFLICT 357 357 T -> A (IN REF. 1).
SQ SEQUENCE 390 AA; 44564 MW; B5F27F986C752CFA CRC64;

Query Match 38.4%; Score 752; DB 1; Length 390;
Best Local Similarity 41.5%; Pred. No. 1.9e-45;
Matches 165; Conservative 81; Mismatches 126; Indels 26; Gaps 10;

QY 1 MASLAANAEEFCNLFREMDNOGNVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
DB 1 MNSLSEANTKFMEDLFOQFRKSEN-NIFSPISITSLALGVLLGAKDNTAQOKRYVLE 59
QY 61 -----NTASGYGNSNSQGL-LOSQLRVSDIASHKQVDLSVNLFLFAKYGFHKQY 114
DB 60 DQVTEITGKAATYUHVDSGNVNHQFQKLTLEFKSYDALELKLANLFEKTYLFLQY 119
QY 115 IECAEKLYDAKVERVDFTNLEDRRNINRWENETHGKIKNYEGEGISSAVMYLVNA 174
DB 120 LDAIKRTYQTSVEVDPAANAPESRRKINSVWESQTEKIKNLIPBENISNTTLVNA 179
QY 175 VYFKGKQSAFTKSETINCHFKSPKSGKAVAMHQRKNLSVIEDPSKILELRYNG- 233
DB 180 IYFGQWEKKFKNEDTKEEKFMPKNTYKSIQWRYQTSFHFASLSDVQAKVLEIPYK 239
QY 234 GINMYVLLPE--NDLSIEENKLTFOUMLEMTNPRRTSKYVEVEFPQFKIEKNYEMQYL 291
DB 240 DLSIVLPLNEIDGLQLEKLTAEKLEMTSLQNMKRETRVDLHLPEKVEESYDKLDL 299
QY 292 RALGLKIDFDESKADLSGIASGRLYISRMHKSYLEVTEGTEAT-----GSNIVE 346

```

Db      300 RTMGWNIJFN-GDADLSCGTMGSGRLSGVLAHAFVEVEEGEAAAATAVVGFGSS--- 355
OY      347 KOLPOST--LFRADHPFLFVIR--KDDILFSGKVCSP 380
Db      356 ---PTSTNEEFCHNPFLEFTRONKNTNSILFGRFSSP 390

RESULT 5
SCC2_HUMAN
ID      SCC2_HUMAN          STANDARD:          PRT;          390 AA.
AC      P48594;
DT      01-FEB-1996 (Rel. 33, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DE      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Squamous cell carcinoma antigen 2 (SCCA-2) (Leupin).
GN      SERPINB4 OR SCCA2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
[1]
RX      MEDLINE-95241462; PubMed-7724531;
RA      Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,
RA      Treter S.D., Hui S.M., Silverman G.A.;
RT      "A serine proteinase inhibitor locus at 18q21.3 contains a tandem
RT      duplication of the human squamous cell carcinoma antigen gene.";
RT      Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
[2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-96013887; PubMed-7589435;
RA      Barnes R.C., Morral D.M.;
RT      "Identification of a novel human serpin gene: cloning sequencing and
RT      expression of leupin.";
RT      FEBS Lett. 373:61-65(1995).
[3]
RN      SEQUENCE FROM N.A.
RX      MEDLINE-21167379; PubMed-11267667;
RA      Hamada K., Shirokawa H., Asano Y., Kihana T., Iwamoto M., Hanakawa Y.,
RA      Hashimoto K., Hirose S., Kyo S., Ito M.;
RT      "Molecular cloning of human squamous cell carcinoma antigen 1 gene and
RT      characterization of its promoter.";
RT      Biochim. Biophys. Acta 1518:124-131(2001).
[4]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Bone marrow;
RA      Strausberg R.;
RL      Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
CC      IMMUNE RESPONSE AGAINST TUMOR CELLS.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.
CC      -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U19576; AAA92602.1; -
DR      EMBL; U19570; AAA92602.1; JOINED.
DR      EMBL; U19571; AAA92602.1; JOINED.
DR      EMBL; U19572; AAA92602.1; JOINED.
DR      EMBL; U19574; AAA92602.1; JOINED.
DR      EMBL; U19575; AAA92602.1; JOINED.
DR      EMBL; U19557; AAA92602.1; JOINED.
DR      EMBL; X89015; CAA61420.1; -
DR      EMBL; AB035089; BAB21525.1; -
DR      EMBL; BC017401; AAH17401.1; -

```

```

DR      HSSP; P01008; 1A7H.
DR      Genew; HGNC:10570; SERPINB4.
DR      MIM; 600518; -.
DR      InterPro; IPR000215; Serpin.
DR      Pfam; PF00079; serpin_1.
DR      SMART; SM00093; SERPIN; 1.
DR      PROSITE; PS00284; SERPIN; 1.
KW      Serpin; Serine protease inhibitor.
FT      ACT_SITE 354 355
SQ      SEQUENCE 390 AA; 44854 MW; 04E213CD892587D5 CRC64;

Query Match      38.2%; Score 748; DB 1; Length 390;
Best Local Similarity 40.8%; Pred. No. 3,6e-45;
Matches 160; Conservative 84; Mismatches 134; Indels 14; Gaps 8;

OY      1 MASIAAANPEFENIFREDDNOGNGNVFESSLSFAALATVRLGADDSLSIDKLLHV 60
DB      1 MNSLSEANRKFMDLEQOFRKSEN-NIFYSTSLGAVLLAKDKNTAQISKVLHF 59
OY      61 -----NTASGYGNSSNSQSG-LQSLKRVFSDINASHKDYLSIYNGLFAEKVGFHKDY 114
DB      60 DQVTEHTTKATYHYHDSRGVNHQFOKLTFFENKSTDAVELKANKLGEKTYQELQDY 119
OY      115 ICAEKLDAKVERVDFTNHLDDTRRNINKWVENETHGKIKVYIGBGCISSAVMYLVNA 174
DB      120 LPAIKKFYOTVSESTDFAVAPESRKRKINSWESQTEKIKMLFPDGTIGDITTLVLYNA 179
OY      175 VYFKGWOSAFKSEFTINCHFKSPQSGKAVAMHMOERKFNLSVIEDPSMKLTLELYNG- 233
DB      180 IIFKGWENKFKKENTKEKFPNKNNTKYSVMQKQYNSFNALLEDVQAKLEIPYKQK 239
OY      234 GIMTYVLPPE--NDLSELENKLTFLQNLMEWTPRRRTSKYVEVFPQPKIEKNYEMKQYL 291
DB      240 DLSMIVLPENEDIGLQKLEKLTJAKIKMWTSLQNMRETCVDLHLPFRFMESSYDLKDTL 299
OY      292 RALGKIDFDESKADSLASGGRLYISMMMKSTIEVTEETATGATGSIYVEKQDPQ 351
DB      300 RTMGWNIJFN-GDADLSCGTMGSHGLSVSLVLAHAFVEVEEGEAAAATAVVGFLSSPS 358
OY      352 ST-LFRADHPFLFVIR--KDDILFSGKVCSP 380
DB      359 TNEEFCHNPFLEFTRONKNTNSILFGRFSSP 390

RESULT 6
ILEU_HORSE
ID      ILEU_HORSE          STANDARD:          PRT;          379 AA.
AC      P05619;
DT      01-NOV-1988 (Rel. 09, Created)
DT      01-FEB-1994 (Rel. 28, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Leucocyte elastase inhibitor (LEI).
GN      SERPINB1 OR ELANH2.
OS      Equus caballus (Horse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
[1]
NX      NCB1_TaxID-9796;
RN      SEQUENCE FROM N.A.
RX      MEDLINE-93319507; PubMed-7687128;
RA      Kordula T., Dublin A., Schooltink H., Koj A., Heinrich P.C.,
RA      Rose-John S.;
RT      "Molecular cloning and expression of an intracellular serpin: an
RT      elastase inhibitor from horse leucocytes.";
RT      Biochem. J. 293:187-193(1993).
[2]
RN      TISSUE=Leukocyte;
RP      SEQUENCE.
RX      MEDLINE-92202200; PubMed-1551869;
RA      Dublin A., Travits J., Enghild J.J., Potempa J.;
RT      "Equine leucocyte elastase inhibitor. Primary structure and
RT      identification as a thymosin-binding protein.";
RT      J. Biol. Chem. 267:6576-6583(1992).

```

RN [3]
 RP SEQUENCE OF 343-362.
 RX MEDLINE=88213423; PubMed=3366785;
 RA Potempa J., Dublin A., Matorek W., Travis J.;
 RT "An elastase inhibitor from equine leukocyte cytosol belongs to the
 RT serpin superfamily. Further characterization and amino acid sequence
 RT of the reactive center.";
 RL J. Biol. Chem. 263:7364-7369(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=92389327; PubMed=1518052;
 RA Baumann U., Bode W., Huber R., Travis J., Potempa J.;
 RT "Crystal structure of cleaved equine leucocyte elastase inhibitor
 RT determined at 1.95-A resolution.";
 RL J. Mol. Biol. 226:1207-1218(1992).
 CC -1- FUNCTION: THIS INHIBITOR IS THOUGHT TO BE INVOLVED IN THE CONTROL
 CC OF INTRACELLULAR PROTEIN TURNOVER. BINDS TO THYMOSIN BETA-4.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M91161; AAA97513.1; -
 DR PIR: A28060; A28060.
 DR PIR: A37276; A37276.
 DR PIR: A42421; A42421.
 DR PIR: S34062; S34062.
 DR PDB: 1HEF; 3I-TAN-94.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 KM Serpin: serine protease inhibitor; 3d-structure.
 FT MOD_RES 1 1 BLOCKED (WITH AN UNKNOWN CHEMICAL GROUP).
 FT ACT_SITE 344 345
 FT HELIX 2 22 REACTIVE BOND.
 FT STRAND 28 30
 FT HELIX 32 44
 FT TURN 45 45
 FT HELIX 48 58
 FT TURN 59 59
 FT HELIX 60 62
 FT TURN 64 65
 FT HELIX 66 77
 FT TURN 78 78
 FT STRAND 85 95
 FT TURN 96 97
 FT HELIX 102 112
 FT STRAND 115 119
 FT TURN 121 123
 FT HELIX 125 139
 FT TURN 140 142
 FT TURN 150 151
 FT TURN 155 156
 FT STRAND 157 171
 FT HELIX 177 179
 FT STRAND 181 186
 FT STRAND 192 209
 FT HELIX 210 212
 FT TURN 213 213
 FT STRAND 214 221
 FT TURN 222 223
 FT STRAND 226 233
 FT HELIX 244 248
 FT TURN 249 249
 FT HELIX 252 259
 FT HELIX 261 263

FT STRAND 265 274
 FT STRAND 276 283
 FT HELIX 285 291
 FT TURN 292 292
 FT HELIX 295 297
 FT TURN 299 301
 FT HELIX 305 308
 FT STRAND 313 326
 FT STRAND 330 343
 FT STRAND 350 353
 FT STRAND 358 364
 FT TURN 365 368
 FT STRAND 369 376
 SO SEQUENCE 379 AA; 43006 MW; 72DED1999371427A CRC64;
 Query Match 37.7%; Score 739.5; DB 1; Length 379;
 Best Local Similarity 38.0%; Pred. No. 1,4e-44;
 Matches 148; Conservative 85; Mismatches 137; Indels 19; Gaps 4;
 QY 1 MASLAANMEFCNLFREDDNNGNVPFSSLSLPAALALVRLGAQDDSLQIDKLLHY 60
 DB 1 MEQLSTANTHFVADLFRALNESDPTGNIFISPLISSALAMTFLGRTAQAQVSALYF 60
 QY 61 NTASGYGNSSNSQSLQSKRVFSDINASHKDYDSIVNGLFAEKVYGFHKDYIECAEK 120
 DB 61 DTV-----EDHSRQSLNADINKGAPYILKANRLYGEKTYNFLADFLASTQK 110
 QY 121 LYDAKVERVDFTHLEDFRRNINKWVENETHGKIKVNIIGEGGISSAVMYLVNAVYFKG 180
 DB 111 MYGAELASVDFQAPARDARKEINEMWCGTEGIPPELLVGMVDNMTKLVLAIVYFKG 170
 QY 181 WQSAFTKSETINCHFSKPCSGKAVAMHOERKFNLSVIEDPMSKLTLELYNG-GINMY 239
 DB 171 WQEKMEKENTRDAPFRNKKVDTVKMKMTOKKFPYNYIEDLCKRVLELPYQKEISMII 230
 QY 240 LIPEN-----DLSEIENKLTFFQNLMEWTNPRRMTSKYVEFPQRIEKNYEMKQYLRA 293
 DB 231 LRPDDLEDESTGEKLEKQVLEKLEMTPEKLYLAENVVHLPRKLEESYDLTSHLAR 290
 QY 294 LGKIDFDESKADLSGIASGGRLYISRMHKSITVEETESTVTAATGSIYVKOLPOST 353
 DB 291 LGVODLFNRGKADLSGMSGARDFEVSIIHKSPEDLNDEGTAAATAGTIMLMLPEE 350
 QY 354 LFRADHPFLFVIRKDI--ILFSGKVSQP 380
 DB 351 NFNADHPFLFIRNPSANILFLGRSSP 379
 RESULT 7
 ID IEU_HUMAN STANDARD: PRT: 379 AA.
 AC P30740;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leukocyte elastase inhibitor (LEI) (Monocyte/neutrophil elastase
 DE inhibitor) (M/NEI) (EI).
 GN SERPINB1 OR ELANH2 OR PI2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92302296; PubMed=1376927;
 RA Remold-O'Donnell E., Chin J., Alberts M.;
 RT "Sequence and molecular characterization of human monocyte/neutrophil
 RT elastase inhibitor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5635-5639(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98296265; PubMed=9630619;
 RA Zeng W., Silverman G.A., Remold-O'Donnell E.;

RT Structure and sequence of human M/NEI (monocyte/neutrophil elastase inhibitor), an Ov-serpin family gene.";
 RL Gene 213:179-187(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervix;
 RA Strausberg R.;
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 178-185; 204-210 AND 364-371.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
 RA Vandekerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 CC -1- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASES ELASTASE, CATHEPSIN G AND PROTEINASE-3.
 CC -1- SUBCELLULAR LOCATION: cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M93056; -; NOT_ANNOTATED_CDS.
 DR EMBL; AF053630; AAC31394.1; -;
 DR EMBL; BC009015; AAH09015.1; -;
 DR PIR; S27383; S27383.
 DR HSSP; P05619; 1HLE.
 DR Aarhus/Ghent-2DPAGE; 4314; IEF.
 DR Genew; HGNC:3311; SERPINB1.
 DR MIM; 130135; -;
 DR InterPro; IPR000215; Serpin.
 DR Pfam; PF00079; serpin.1.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 344 345 REACTIVE_BOND (BY SIMILARITY).
 SQ SEQUENCE 379 AA; 42741 MW; BAAE08DFC8CD3 CRC64;

 Query Match 36.9%; Score 722.5; DB 1; Length 379;
 Best Local Similarity 38.6%; Pred. No. 2.1e+43;
 Matches 150; Conservative 74; Mismatches 146; Indels 19; Gaps 4;

 QY 1 MASLAANAEFCNFLEFREMDDNGNGNFFSSISLFLAALVRLGAQDDSLQIDKILHY 60
 DB 1 MQLSSANTRFALDLFLALSENENNPAGNITIPSSISAAAMVFLGRGTAQDLSTFRF 60
 QY 61 NNASGTGNSNSGSLQSLKRFSDINASHKDYLSYNGLEPAEKVYFPHKDYLECAEK 120
 DB 61 NTV-----EEVHSRFOSIAMDINRKGASYILKLANRLYGEKTYNLFDEFLVSTOK 110
 QY 121 LVDKAVRYDTNHLDETRRNINKWNETHGKIKNVIGEGGSSSAVAVLVNAVYFKK 180
 DB 111 TYGADIASYFQHASHEADAKKTINQWKGTEGTEIPPLASGAWDNFKTKLVNALVYKGN 170
 QY 181 MOSAFETSETINCHFKSPKSGKAVAMHQRKFNLSVIEDSMKLTLELYNG-GINMYV 239
 DB 171 WKDFEKKEATNAPFLNKKDKRYKMMYQKKKFAAGTYEDLKCRLLELPYGEELSMYI 230
 QY 240 LIPEN-----DLSEIENKLTQNTMEWTNPRMTSKYVEVFPQKIEKNTEMOYLRA 293
 DB 231 LIPDIEDESTGLKTEELTLEKLTLEMTKPEMLDIEVNAVSLPRKIEESYTLNSDLAR 290
 QY 294 LGLKDFEDSKADLSGIASGRLYISRMHKSIVYETEGCTATATGSIYKOLPOST 353
 DB 291 LGVODLENSKADLSGMSGARDIFISKIVKSFVEVNEGTEBAATAATAGIATFCMLPBE 350

QY 354 LFRADHPFLFVIRKDD--ILFSGKVS 380
 DB 351 NFRADHPFLFIRHNSGSLFLGRSSP 379

 RESULT 8
 ID SB13_HUMAN STANDARD; PRT; 391 AA.
 AC Q9UIV8; Q9UBW1; Q9UGO; Q9HCX1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hurpin (HaCap UV-repressible serpin) (Protease inhibitor 13)
 DE (Headpin) (Serpin B13).
 GN SERPINB13 OR P113.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Skin;
 RX MEDLINE=9443898; PubMed=10512713;
 RA Abts H.F., Weiss T., Mirmohammadsadegh A., Koehler K., Michel G.,
 RA Ruzicka T.;
 RT "Cloning and characterization of hurpin (Protease inhibitor 13): a new skin specific, UV-repressible serine proteinase inhibitor of the ovalbumin serpin family.";
 RL J. Mol. Biol. 293:29-39(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=9458661; PubMed=10527881;
 RA Spring P., Nakashima T., Frederick M., Henderson Y., Clayman G.;
 RT "Identification and cDNA cloning of headpin, a novel differentially expressed serpin that maps to chromosome 18q.";
 RL Biochem. Biophys. Res. Commun. 264:299-304(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20461850; PubMed=11004515;
 RA Nakashima T., Pak S.C., Silverman G.A., Spring P.M., Frederick M.J.,
 RA Clayman G.L.;
 RT "Genomic cloning, mapping, structure and promoter analysis of HEADPIN, a serpin which is down-regulated in head and neck cancer cells.";
 RL Biochim. Biophys. Acta 1492:441-446(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Placenta;
 RX MEDLINE=21213570; PubMed=11313015;
 RA Abts H.F., Weiss T., Scheuring S., Scott F.L., Irving J.A., Michel G.,
 RA Bird P.L., Ruzicka T.;
 RT "Sequence, organization, chromosomal localization and alternative splicing of the human serine protease inhibitor gene hurpin (P113), which is up-regulated in psoriasis.";
 RL DNA Cell Biol. 20:123-131(2001).
 CC -1- FUNCTION: May play a role in the proliferation or differentiation of keratinocytes.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SKIN-SPECIFIC.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AJ001696; CAA04935.2; -;
 DR EMBL; AJ001697; CAA04936.2; -;

```

DR EMBL: AJ001698; CAA04937.1; -.
DR EMBL: AF169949; AAD5765.1; -.
DR EMBL: AF216854; AAF72879.1; -.
DR EMBL: A278717; CAC03569.1; -.
DR HSSP: P05120; 1BY7.
DR Genew; HGNC:8944; SERPINB13.
DR MIM; 604445; -.
DR InterPro: IPR000240; Maspin.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR PRINTS; PR00676; MASPIN.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
DR PROSITE; Serpine protease inhibitor: Alternative splicing.
KM ACT_SITE 356 357 REACTIVE_BOND (BY SIMILARITY).
FT VARSPIC 206 257 MISSING (IN ISOPFORM 2).
FT CONFLICT 8 8 S -> N (IN REF. 4).
FT CONFLICT 75 75 MISSING (IN REF. 1; CAA04937).
FT CONFLICT 293 293 G -> S (IN REF. 2).
FT CONFLICT 297 297 E -> Q (IN REF. 1; CAA04937).
SQ SEQUENCE 391 AA; 44276 MW; 2CA8858D4BC2B09 CRC64;

Query Match 36.5%; Score 714.5; DB 1; Length 391;
Best Local Similarity 38.0%; Pred. No. 7.9e-43;
Matches 152; Conservative 77; Mismatches 142; Indels 29; Gaps 7;

QY 1 MASLAANAAMAFECNLFREMDNONGNNGVNFSSLSLFAALALVLRGAODDSLQIDKLHY 60
DB 1 MDSLGAAYSTRGLGDLFEEL-KKTNDGNIFSPVGLTALIGVLLGTRGATASOLEEYFH- 58
QY 61 NTASGVGNSSNSOSGLKRVFSDINASHK-----DYDSLIVNGLPFAEK 106
DB 59 -----SEKTKSSRIKAKEKEVIENTEAHOOPOKFLTISKLTNDYELINTNRLEGEK 112
QY 107 VYGFHMDYIECAEKLYDAKVERVDFTNHLDETNRNINKVVENETHGKIKVIGEGGSISS 166
DB 113 TYLFQKYLDYVEKKYHASLEPVDVNAADESRKKISVYESKTEKIKDLFPDGSISSS 172
QY 167 AVAVLVNAVYFKGKMOAFTKSEINCHFKSPKCSGKAVAMMOERKFNLSVIEDPKMKI 226
DB 173 TKLVLVNAVYFKGOMDREFFKENTKEKFMNKSSTKSVOMQOSHSFTFEDLOAKI 232
QY 227 LEDRY-NGGINMYVLLPNDLSEIE--NKLTFQNLMEWTPRRTSKYVEVFPOPKIE 282
DB 233 LGIPYKNDLSMFLLP-NDIDGLEKIDIKISPEKLEWTPSPGMEERKKNVNLHPREVE 291
QY 283 KNEEMKOYLALGLKIDFDSKADLSGIASGRLYISRMHKSYLETEEGTEATATGS 342
DB 292 DGYDLNAVLAAMGMDAFSEHKADYSGMSSGSLYAKFLHSSFVAVTEEGTEAATGI 351
QY 343 NIVEKQLOPSTLFRADHPFLFYIR--KDIILFSGKVSPP 380
DB 352 GFTVTSAPGHENYCHNHPFLFTRHNESNLSLFFGRSSP 391

RESULT 9
OVARY_CHICK STANDARD; PRT; 388 AA.
AC P01014;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Gene y protein (Ovalbumin-related).
GN y
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauilia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83014329; Pubmed=7122240;
RA Hellig R., Muraskovsky R., Kloepper C., Mandel J.L.;

```

```

RT "The ovalbumin gene family: complete sequence and structure of the y
RT gene";
RL Nucleic Acids Res. 10:4363-4382(1982).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: J00922; AAA68882.1; -.
DR PIR: A01244; DYCH.
DR HSSP: P01012; 10VA.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin.1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin.
KM ACT_SITE 353 354 REACTIVE_BOND_HOMOLOG.
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 388 AA; 43772 MW; 2AF17BF150F461 CRC64;

Query Match 36.3%; Score 711; DB 1; Length 388;
Best Local Similarity 39.6%; Pred. No. 1.4e-42;
Matches 155; Conservative 75; Mismatches 147; Indels 14; Gaps 7;

QY 1 MASLAANAAMAFECNLFREMDNONGNNGVNFSSLSLFAALALVLRGAODDSLQIDKLHY 60
DB 1 MDSISVYNAKFCEDVFNEMKVHVNENILICPSILTALMAYLVLRGAGNSESQMKVLYF 60
QY 61 NTASGVGNSSNSOSG----LOSOLKRVFSDINASHKDYDSLIVNGLPFAEKVYGFHMDYIE 116
DB 61 DSITGAGSTIDSCGSEYVHNLFKELSETTRNATYSLIEDKLYVDFTFSLPYLS 120
QY 117 CAEKLYDAKVERVDFTNHLDETNRNINKVVENETHGKIKVIGEGGSISSAVLVNAVY 176
DB 121 CARFYGVEEYVFNKAAAEARQLINSVWEKETNGOIKDLVSSSIDFTWVFINTIY 180
QY 177 FKGMOSAFTKSEINCHFKSPKCSGKAVAMMOERKFNLSVIEDPKMKILERY-NGGI 235
DB 181 FKGIWKAETEDTREMPEFMTKEESKPVOMCMNNSFNATLPAEKMKITLLEPYASGDL 240
QY 236 NMVYLPE--NDLSEIENKLTFOQLMEWTPRRTSKYVEVFPOPKIEKNYEMKOYLRA 293
DB 241 SMVLPLPDEVSGLEIKETINFDKLRMTSTNMAKMSMKVYLPKMKIEKYNLTSLMA 300
QY 294 LGLKIDFDSKADLSGIASGRLYISRMHKSYLETEEGTEATATGS--NIVEKQLOPQ 351
DB 301 LGMWDLFSRS-ANLTGSSVDNLMISDAVGEVEEVEEGTEATGSAIGNI--KHSLE 357
QY 352 STLEFRADHPFLFYIRKD--DIILFSGKVSPP 380
DB 358 LEERADHPFLFYIRYNPTNALILFGKRWSP 388

RESULT 10
PT16_BOVIN STANDARD; PRT; 378 AA.
AC 002739;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine proteinase inhibitor B-43.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.

```


OX NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97165893; PubMed=9013786;
 RA Nakaya N., Nishihori M., Kawabata M., Saeki K.;
 RT "Cloning of a serine proteinase inhibitor from bovine brain:
 RT expression in the brain and characterization of its target
 RT proteinases.";
 RL Brain Res. Mol. Brain Res. 42:293-300(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF SERINE PROTEINASES
 CC PRESENT IN THE BRAIN OR EXTRASARIAL FROM THE BLOOD.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D55670; BAA19875.1; -
 DR HSSP: P05120; 1BY7.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 DR Serpin; Serine protease inhibitor.
 FT ACT_SITE 343
 FT 344
 SQ SEQUENCE 378 AA; 42560 MW; 664F499CCCE263A CRC64;

 Query Match 35.8%; Score 702; DB 1; Length 378;
 Best Local Similarity 39.4%; Pred. No. 5.6e-42;
 Matches 152; Conservative 83; Mismatches 137; Indels 14; Gaps 7;

 OY 1 MASLAANAEEFCNLFREDDNDGNG-NYFFSLSLFALALVRLGADDSLSQIDKLH 59
 DB 1 MDALSAANGTFLTLTKL-GEQNSKNVFISPLSISALAVLLGAKNTAAQMGQTLS 58
 OY 60 VNTASGYGNSNSQSGSLQKRVFSDINASHKVDLSVNGLFAEKVYGFHKDYECAE 119
 DB 59 LKSSGSG--EDVHQGFQNL-----SEVNRDPTQYLLRTANRLFGKTYDFLSSFDSC 112
 OY 120 KLYDAKVERVDFTNHNEDFRNRINKVNEETHGKIKINVGEGISSAAMVAVNAVYFG 179
 DB 113 KYQAMERLDVVSATQSRKHNTWAEKTBKTRDLPANSVNPTRVLVNAVATYFGK 172
 OY 180 KMOSAFETKSETINCHFKSPKCGKAVAMHQRKFNLSYIEDPSMKILELRNG-GINNY 238
 DB 173 NMDTQFNKEETDEPRFVRSKNEKPYQMFKSTCKITYIGELISQIDLVLPVGGELNAV 232
 OY 233 VLLP--ENDLSEIKLTFQNLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLRAQL 296
 DB 233 ILLPESDLDLNTVEALYEFKFIAMTKPDVMDDEEVEVFLPFTLEESYDMEFLOELGM 292
 OY 297 KLIPEESKADLSGIASGRLYISRMMKSYIEVTEGTEATATGNSINYEKOLPOSTLFR 356
 DB 293 TDAFEETRADFSGMSGRGLHLSKVHNSFVEVTEGTEAATAAGVAVMMKCLVAVPRPN 352
 OY 357 ADHPFLFVIR--KDIILFSGKVS 380
 DB 353 ANHPFLFVIRHSGKTAALLFCGRFCSP 378

 RESULT 11
 ID ILEU_PIG STANDARD: PRT; 378 AA.
 AC P80229;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leukocyte elastase inhibitor (LEI) (Leukocyte neutral proteinase
 DE inhibitor) (LNP1).
 GN SERPINB1 OR ELANH2.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN (1)
 RP SEQUENCE.
 RC TISSUE=Leukocyte;
 RX MEDLINE=94039085; PubMed=7901009;
 RA Teschauer W.F., Mentle R., Sommerhoff C.P.;
 RT "Primary structure of a porcine leukocyte serpin.";
 RL Eur. J. Biochem. 217:519-526(1993).
 CC -1- FUNCTION: REGULATES THE ACTIVITY OF THE NEUTROPHIL PROTEASE
 CC ELASTASE, CATHESPIN G AND PROTEINASE-3 (By similarity).
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 DR HSSP: P05619; 1HLE.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 DR Serpin; Serine protease inhibitor.
 FT MOD_RES 1
 FT DISULFID 80 80
 FT 343 344 INTERCHAIN (PROBABLE).
 FT ACT_SITE 343 344 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 378 AA; 42512 MW; 3CA7C65F11E508CA CRC64;

 Query Match 35.5%; Score 696; DB 1; Length 378;
 Best Local Similarity 37.5%; Pred. No. 1.5e-41;
 Matches 146; Conservative 82; Mismatches 141; Indels 20; Gaps 5;

 OY 1 MASLAANAEEFCNLFREDDNDGNGNFFSSLFALALVRLGADDSLSQIDKLH 60
 DB 1 MDLSAANTRFALDLFRALNESNPAGNIFISPSISALAMILLGRTGTEAOMKALHF 60
 OY 61 NTASGYGNSNSQSGSLQKRVFSDINASHKVDLSVNGLFAEKVYGFHKDYECAE 120
 DB 61 DTV-----KDIHSFQSLANDINKGASLYLKLANKRLEFGKTYHFLPEFLASTQK 110
 OY 121 LYDAKVERVDFTNHNEDFRNRINKVNEETHGKIKINVGEGISSAAMVAVNAVYFGK 180
 DB 111 TYGAELASVDPLRASBEAKAINEMWKEQTEBKIPELLASGVDSATKLVNAVATYFGK 170
 OY 181 WOSATKSETINCHFKSPKCGKAVAMHQRKFNLSYIEDPSMKILELRNG-GINNY 239
 DB 171 WQEKMTETATKDAAPRLNKKDSKTYKMMYQKRRFFGYIKELKCHVLELPYOGKDLNVI 230
 OY 240 LLPEN-----DLSEIKLTFQNLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLRA 293
 DB 231 LLPDSTEDSTGLRKIEQHTLEKLNEMTKPRLNLELVNAVLPFRLESDLVNAPLAR 290
 OY 294 IGLKIDFESKADLSGIASGRLYISRMMKSYIEVTEGTEATATGNSINYEKOLPOST 353
 DB 291 LGVQDLFG--SRADLGMSEARDLFTSKYVHKAFVEVNEEGTEAATAAGVAVFAMLPDE 349
 OY 354 LFRADHPFLFVIRK--DIILFSGKVS 380
 DB 350 DFIADHPFLFVIRHNPSSNILELRISSP 378

 RESULT 12
 ID PAI2_HUMAN STANDARD: PRT; 415 AA.
 AC P05120; Q96E96;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE plasminogen activator inhibitor-2 precursor (PAI-2) (Placental
 DE plasminogen activator inhibitor) (Monocytic Arg-serpin) (Urokinase

DE inhibitor).
GN SERPINB2 OR PAI2 OR PLANH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE=87137674; PubMed=3029122;
RA Ye R.D., Mun T.-Z., Sadler J.E.;
RT "CDN cloning and expression in Escherichia coli of a plasminogen
RT activator inhibitor from human placenta.";
RL J. Biol. Chem. 262:3718-3725(1987).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=88142852; PubMed=3325828;
RA Schleuning W.-D., Medcalf R.L., Hession C., Rothenbuhler R.,
RA Shaw A., Krulthof E.K.O.;
RT "Plasminogen activator inhibitor 2: regulation of gene transcription
RT during phorbol ester-mediated differentiation of U-937 human
RT histiocytic lymphoma cells.";
RL Mol. Cell. Biol. 7:4564-4567(1987).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=89174589; PubMed=2494165;
RA Ye R.D., Ahern S.M., Le Beau M.M., Lebo R.V., Sadler J.E.;
RT "Structure of the gene for human plasminogen activator inhibitor-2.
RT The nearest mammalian homologue of chicken ovalbumin.";
RL J. Biol. Chem. 264:5495-5502(1989).
RN (4)
RP SEQUENCE FROM N.A.
RC TISSUE-Monocytes;
RX MEDLINE=87252928; PubMed=3496414;
RA Webb A.C., Collins K.L., Snyder S.F., Alexander S.J.,
RA Rosenwasser L.J., Eddy R.L., Shows T.B., Auron P.E.;
RT "Human monocyte Arg-Serpin cDNA. Sequence, chromosomal assignment,
RT and homology to plasminogen activator-inhibitor.";
RL J. Exp. Med. 166:77-94(1987).
RN (5)
RP SEQUENCE FROM N.A.
RC TISSUE-Monocytes;
RX MEDLINE=88125032; PubMed=3257578;
RA Antalis T.M., Clark M.A., Barnes T., Lehtbach P.R., Devine P.L.,
RA Schevov G., Goss N.H., Stephens R.W., Tolstoshev P.;
RT "Cloning and expression of a cDNA coding for a human monocyte-derived
RT plasminogen activator inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:985-989(1988).
RN (6)
RP SEQUENCE FROM N.A.
RX MEDLINE=90152678; PubMed=2303256;
RA Samia J.A., Alexander S.J., Horton K.W., Auron P.E., Byers M.G.,
RA Shows T.B., Jr., Webb A.C.;
RT "Chromosomal organization and localization of the human urokinase
RT inhibitor gene: perfect structural conservation with ovalbumin.";
RL Genomics 6:159-167(1990).
RN (7)
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN (8)
RP SEQUENCE OF 12-17, 103-108 AND 314-321.
RX MEDLINE=93162043; PubMed=1286667;
RA Resmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandeckerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
RN (9)
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=9918112; PubMed=10368272;
RA Harrop S.J., Jankova L., Coles M., Jardine D., Whittaker J.S.,

RA Gould A.R., Meister A., King G.C., Mabbutt B.C., Curmi P.M.G.;
RT "The crystal structure of plasminogen activator inhibitor 2 at 2.0-A
RT resolution: implications for serpin function.";
RL Structure 7:43-54(1999).
CC CC THE MONOCYTE DERIVED PAI-2 IS DISTINCT FROM THE ENDOTHELIAL
CC CELL-DERIVED PAI-1.
CC -1- SUBCELLULAR LOCATION: CYTOSOL; PLASMA MEMBRANE; EXTRACELLULAR.
CC -1- PFM: THE SIGNAL SEQUENCE IS NOT CLEAVED.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02685; AAA36413.1; -;
DR EMBL: M18082; AAA60006.1; -;
DR EMBL: Y00630; CAA68666.1; -;
DR EMBL: M24657; AAA60348.1; -;
DR EMBL: M24651; AAA60348.1; JOINED.
DR EMBL: M24652; AAA60348.1; JOINED.
DR EMBL: M24653; AAA60348.1; JOINED.
DR EMBL: M24654; AAA60348.1; JOINED.
DR EMBL: M24655; AAA60348.1; JOINED.
DR EMBL: M24656; AAA60348.1; JOINED.
DR EMBL: J03603; AAA60004.1; -;
DR EMBL: M31551; AAA36797.1; -;
DR EMBL: M31547; AAA36797.1; JOINED.
DR EMBL: M31548; AAA36797.1; JOINED.
DR EMBL: M31549; AAA36797.1; JOINED.
DR EMBL: M31550; AAA36797.1; JOINED.
DR EMBL: BC012609; AAH12609.1; -;
DR EMBL: A21238; CAA01535.1; -;
DR EMBL: A21254; CAA01539.1; -;
DR PIR: A26553; A26553.
DR PIR: A29362; A29362.
DR PIR: A32853; A32853.
DR PIR: A31366; A31366.
DR PDB: 1BY7; 24-OCT-99.
DR Aarhus/Chent-2DPAGE: 6314; IEF.
DR Genew; HGNC:8584; SERPINB2.
DR MIM; 173390; -;
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin.1.
DR SMART; SM00093; SERPIN.1.
DR PROSITE; PS00284; SERPIN.1.
KW Serpin; Serine protease inhibitor; Plasma; Plasminogen activation;
KW Glycoprotein; Signal; 3D-structure; Polymorphism.
FT SIGNAL 1 ?
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT ACT_SITE 380 381 REACTIVE BOND.
FT VARIANT 120 120 N -> D (IN DBSNP: 6098).
FT VARIANT 404 404 /FTID=VAR.011743.
FT VARIANT 413 413 N -> K (IN DBSNP: 6103).
FT VARIANT 413 413 /FTID=VAR.011744.
FT VARIANT 413 413 S -> C (IN DBSNP: 6104).
FT CONFLICT 170 170 /FTID=VAR.011745.
FT CONFLICT 170 170 N -> Y (IN REF. 7).
SQ SEQUENCE 415 AA; 46596 MW; 10DFAB5A4B1246FF CRC64;
Query Match 33.9%; Score 664.5; DB 1; Length 415;
Best Local Similarity 34.7%; Pred. No. 2.6e-39;
Matches 144; Conservative 83; Mismatches 153; Indels 35; Gaps 5;
Qy 1 MASIAANAEFCENFREMNDONGNCFSSSLFPAALVALVRLGAQDSDLQIDRLHLY 60
Db 1 MEDLCVANTLFALNLFHLAKASPTQNLFLSPWSISSTYAMVYWGSRGSTEQMAKVLQF 60

QY 61 N-----TASGY-----GN-----SSNSOGSOLKRVFSDINASHKD 93
 DB 61 NEGANAAVTPMTPENTSCGFMQIQKGSYPAIILOAQADKHHSSFRSLSSAINSTON 120
 QY 94 YDLSTVNGFAEKYGVGFHHDYIECAEKLKDYAKVERDFNHLNEDTRRNINKVVENETHGK 153
 DB 121 YLLESVNKLFGESASFEIREYIRLCOKYSSSEPOAVDFLECAEARKKINSWVKCTQTK 180
 QY 154 IKNVIGEGLSSAVVAVLVNAVYFKGMQSAFTKSEFTINCHFPKSPKCSGKAVAMHMQEKK 213
 DB 181 IENLPEGSVDGDTKRAVLVNAVYFKGMQSAFTKSEFTINCHFPKSPKCSGKAVAMHMQEKK 240
 QY 214 FNLVSEIDPSMKILEIRYNGINMYVLPE-----NDLSEIENKLTFOINLMEWTPRRM 267
 DB 241 LKIGYEDDKAOILELPYAGDVSMFLLPDEIADVSTGLLESEITTYDKLNKMTSKDKM 300
 QY 266 TSKYVEVFEPKRIENKEMKQYLRALGKIDFDESKADLSGASGRLYISMMHKSIT 327
 DB 301 AEDEVEVYIPORLEHVELRSLRSMGMDAFNKRANFGSGSERNDLFLSEVFHQAVY 360
 QY 328 EYTEEGTEATNGSNIVEKQLPOSTLFRADHPFLFVI--RKDDILFGSKVSCP 380
 DB 361 DVNEEGTEAAGTGGVMTGRTHGGRQGVADHPFLFLMHKTKITNCILFGGRSSP 415

RESULT 13

ID SPB8_HUMAN STANDARD: PRT: 374 AA.
 AC P50452:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytoplasmic antiprotease 2 (CAP2) (CAP-2) (Protease inhibitor 8)
 DE (Serpin B8).
 GN SERPINB8 OR P18.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96102039; PubMed=8530382;
 RA Sprecher C.A., Morgenstern K.A., Mathews S., Dahlen J.R.,
 RA Schrader S.K., Foster D.C., Kistiel W.;
 RT "Molecular cloning, expression, and partial characterization of two
 RT novel members of the ovalbumin family of serine protease
 RT inhibitors";
 RL J. Biol. Chem. 270:29854-29861(1995).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L40377; AAC1939.1; -
 DR HSSP: P05120; 1B77.
 DR Genew: HGNC:8952; SERPINB8.
 DR MIM: 601697; -
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 KW Serpin; Serine protease inhibitor.
 FT ACT SITE 339 340 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 374 AA. 42786 MW. E855A03519AC60E CRC64;

Query Match 33.8%; Score 662; DB 1; Length 374;
 Best Local Similarity 37.4%; Pred. No. 3,46-39;
 Matches 145; Conservative 79; Mismatches 142; Indels 22; Gaps 8;

QY 1 MASLAANAEPFCNLFREDDNNGNNGVFFSSLSLFAALALVRLGAODSLSDIDKLHV 60
 DB 1 MDLCEANCTFAISLEKILIGEDNSRNVEFSPMSISALAMVFGKSGSTAOMSQALCL 60
 QY 61 NTASGNGNSNSOGSOLKRVFSDINSHHDYDIYNGFAEKYGVGFHHDYIECAEK 120
 DB 61 Y-----KGGDHLRGQSIL-----SEVNTGTQYILRTANRFLGETCTCFLEDFKRYCK 110
 QY 121 LYDAKVERDFNHLNEDTRRNINKVVENETHGKIKKNGVEGSSAVVAVLVNAVYFKGK 180
 DB 111 FYQAELEELSFADIEECHKHINDVAEETEKISEVLDAGTVDPDKLVNAVYFKGK 170
 QY 181 WQSAFTKSEFTINCHFPKSPKCSGKAVAMHMQEKKFNLSYEDSMKLEIRY-NGGINMYV 239
 DB 171 WNEQDFRKYTKRMLEFTE-ERKTYQMPKREKFGYADVHTVLELPYEEELSWYI 229
 QY 240 LLPEN--DLSEIENKLTFOINLMEWTPRRMTSKYVEVFEPKRIENKEMKQYLRALGK 297
 DB 230 LLPDQNTDLAVVEKALTYEKFKAWTNSEKLTYSKVQVFLPKLLESTYDLERLGM 289
 QY 298 DIFDESKADLSGASGRLYISMMHKSIEYTEEGTEATNG--SNIVEKQLPOSTL 354
 DB 290 DAFDEKADFGSGMSTERNVPLSKVAHCKEVEVNEEGTEAATAAVAVRNSCRSMRPR--- 346
 QY 355 FRADHPFLFVIR--KDDILFGSKVSCP 380
 DB 347 FCADHPFLFVIRRHKTNCILFGGRSSP 374

RESULT 14

ID SB10_HUMAN STANDARD: PRT: 397 AA.
 AC P48595;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Bomanin (Protease inhibitor 10) (Serpin B10).
 DE SERPINB10 OR P110.
 GN SERPINB10 OR P110.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=bone marrow;
 RX MEDLINE=96070759; PubMed=7592909;
 RA Riewald M., Schlegel R.R.;
 RT "Molecular cloning of bomanin (protease inhibitor 10), a novel human
 RT serpin that is expressed specifically in the bone marrow";
 RL J. Biol. Chem. 270:26754-26757(1995).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF PROTEASE
 CC ACTIVITIES DURING HEMATOPOIESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE BONE MARROW.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U35459; AAC50282.1; -
 DR HSSP: P05619; 1HLE.
 DR Genew: HGNC:8942; SERPINB10.
 DR MIM: 602058; -
 DR InterPro: IPR000215; Serpin.

DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 362 363
 SQ SEQUENCE 397 AA: 45402 MW: 8CE0124686715ADF CRC64;

Query Match 33.6%; Score 658.5; DB 1; Length 397;
 Best Local Similarity 34.9%; Pred. No. 6,5e-39;
 Matches 139; Conservative 89; Mismatches 151; Indels 19; Gaps 5;

OY 1 MASLAANAFECFNLREMDNNGCNVFFSSLSLEPAALALVRLGADDSLSQIDKLHY 60
 D 1 MSLATSIHQFALESLKLAESAOGKNIFSSWSISTLIYVLGAKGTAAQAOVLQF 60
 OY 61 NIASGCG-----NSNSGSGLSQSLKRPFSQIMASHKRYDLSIVGLAEKY 107
 D 61 NRDOGVKCDPESEKRRKMEFNLSNSEE-IHSDFOTLISELKLKPDNDYLLKTANAIVGEKT 119
 OY 108 YGFHKDYICAEKLYDAKVERVDFTNHLIEDTRINKVVENETHGKIKNYIGGGISSA 167
 D 120 YAFHNKYLIEDMKYFPAEPQPVAFASQDIRKQINSWVEROTEGKIONLLPDDSDST 179
 OY 168 VAVLVNAVYFKGKMQSAFTKSETINCHFKSPKSGRAVAMHOREKFNLSVIDEPMKIL 227
 D 180 RMLVALVLFKGIWEHQFLVQNTTEKPRINETTSKPYQMMFKKHLFIHEIKPRAVGL 239
 OY 228 ELRYNG-GINMYVLDE--NDLSEINKLTFQNLMTNPRRTSYVEVEYFPOFKIEKN 284
 D 240 QLYKSRDLSLLLPEDINGLEOLEKATYEKLNMTSADNMELEYEVOLHLKPKFLIEDS 299
 OY 285 YEMKQYLRALGLKIDESKADLSIASGRLYISRMHKSRYEVEEGETAATGASNI 344
 D 300 YDLSKTLSSKMSDASOSKADSGMSANLFLSNVFAFVEINDEGTEAAGSGSEI 359
 OY 345 VEKQLPSTLFRADHPFLFVIR--KDDILFSGKVCSP 380
 D 360 DIRIRVPSIEFNANHPFLFIRHNKNTILFYGRLLSP 397

RESULT 15
 OVAL_CHICK
 ID OVAL_CHICK STANDARD; PRT; 385 AA.
 AC P01012;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 01, Last annotation update)
 DE Ovalbumin (Plakalbumin) (Allergen Gal d 2) (Gal d II).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82069038; PubMed=6272839;
 RA Woo S.L.C., Beattie W.G., Catterall J.F., Dugaiczky A., Staden R.,
 RA Brownlee G.G., O'Malley B.W.;
 RT "Complete nucleotide sequence of the chicken chromosomal ovalbumin
 RT gene and its biological significance.";
 RL Biochemistry 20:6437-6446(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=78199842; PubMed=661981;
 RA McReynolds L., O'Malley B.W., Nisbet A.D., Fothergill J.E., Givol D.,
 RA Fields S., Robertson M., Brownlee G.G.;
 RT "Sequence of chicken ovalbumin mRNA.";
 RL Nature 273:723-728(1978).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=79010682; PubMed=692731;
 RA Catterall J.F., O'Malley B.W., Robertson M.A., Staden R.,
 RA Tanaka Y., Brownlee G.G.;

RT "Nucleotide sequence homology at 12 Intron-exon junctions in the
 RT chick ovalbumin gene.";
 RL Nature 275:510-513(1978).
 RN [4]
 RP SEQUENCE OF 1-35.
 RX MEDLINE=7816057; PubMed=272676;
 RA Palmiter R.D., Gagnon J., Maish K.A.;
 RT "Ovalbumin: a secreted protein without a transient hydrophobic leader
 RT sequence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:94-98(1978).
 RN [5]
 RP SEQUENCE OF 1-16.
 RX MEDLINE=79186958; PubMed=751625;
 RA Thompson E.O.P., Fisher W.K.;
 RT "A correction and extension of the acetylated amino terminal sequence
 RT of ovalbumin.";
 RL Aust. J. Biol. Sci. 31:443-446(1978).
 RN [6]
 RP SEQUENCE OF 5-16; 29-35; 60-78; 115-123; 366-373 AND 379-385.
 RX MEDLINE=79186957; PubMed=751624;
 RA Thompson E.O.P., Fisher W.K.;
 RT "Amino acid sequences containing half-cysteine residues in ovalbumin.";
 RL Aust. J. Biol. Sci. 31:433-442(1978).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=90370102; PubMed=2395463;
 RA Stein P.E., Leslie A.G.W., Finch J.T., Turnell W.G., McLaughlin P.J.,
 RA Carrell R.W.;
 RT "Crystal structure of ovalbumin as a model for the reactive centre of
 RT serpins.";
 RL Nature 347:99-102(1990).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=92046044; PubMed=1942038;
 RA Stein P.E., Leslie A.G.W., Finch J.T., Carrell R.W.;
 RT "Crystal structure of uncleaved ovalbumin at 1.95-A resolution.";
 RL J. Mol. Biol. 221:941-959(1991).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=90278960; PubMed=2352279;
 RA Wright H.T., Qian H.X., Huber R.;
 RT "Crystal structure of plakalbumin, a proteolytically nicked form of
 RT ovalbumin. Its relationship to the structure of cleaved alpha-1-
 RT proteinase inhibitor.";
 RL J. Mol. Biol. 213:513-528(1990).
 RN [10]
 RP REVIEW.
 RX MEDLINE=21312433; PubMed=11419711;
 RA Huntington J.A., Stein P.E.;
 RT "Structure and properties of ovalbumin.";
 RL J. Chromatogr. B 756:189-198(2001).
 CC -1- FUNCTION: Not known.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Major protein of egg white.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -1- DATABASE: NAME=worthington biochem manual; OW-SERPIN SUBFAMILY.
 CC WWW="http://www.worthington-biochem.com/manual/OA.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J00895; AAB59956.1; -;
 DR EMBL: V00438; CAA23716.1; -;
 DR EMBL: V00383; CAA23682.1; -;
 DR EMBL: M34352; AAA48998.1; -;
 DR EMBL: M34346; AAA48998.1; JOINED.
 DR EMBL: M34347; AAA48998.1; JOINED.
 DR EMBL: M34348; AAA48998.1; JOINED.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 165.904 Seconds
(without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-34
Perfect score: 1959
Sequence: 1 MASLAANAFCFNLFRMD.....FLFVIRKDDIILFSGKVSQP 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_bacteriophage:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1505	76.8	380	11 09D695	09D695 mus musculus
2	1488	76.0	380	11 09D695	09D695 mus musculus
3	800	40.8	388	11 091212	091212 mus musculus
4	798	40.7	388	11 09COV3	09COV3 mus musculus
5	763.5	39.0	423	11 09D7P9	09D7P9 mus musculus
6	708.5	36.2	379	11 09D154	09D154 mus musculus
7	707	36.1	382	11 08VHP7	08VHP7 mus musculus
8	700.5	35.8	379	11 08VHP7	08VHP7 mus musculus
9	694.5	35.5	369	4 09BYF7	09BYF7 mus musculus
10	686.5	35.0	387	11 09D105	09D105 mus musculus
11	684.5	34.9	377	11 09D804	09D804 mus musculus
12	684.5	34.9	385	11 08VH00	08VH00 mus musculus
13	682	34.8	388	11 09Z2G2	09Z2G2 mus musculus
14	663	33.8	386	13 073860	073860 meleagris g
15	658	33.6	386	11 09D1E7	09D1E7 mus musculus
16	655	33.4	338	4 09BYF8	09BYF8 homo sapien

17	645	32.9	374	11 008797	008797 mus musculus
18	624.5	31.9	377	11 09D6V6	09D6V6 mus musculus
19	619	31.6	374	11 008800	008800 mus musculus
20	606.5	31.0	377	11 008806	008806 mus musculus
21	601.5	30.7	379	11 09D058	09D058 mus musculus
22	589.5	30.1	377	11 08VH01	08VH01 mus musculus
23	585.5	29.9	453	13 08VVS2	08VVS2 struthio ca
24	583.5	29.8	377	11 09D827	09D827 mus musculus
25	579.5	29.6	456	13 08VVS0	08VVS0 xenopus lae
26	573.5	29.3	453	13 08VVS1	08VVS1 cheilydra se
27	570.5	29.1	452	13 09PTA8	09PTA8 salmo salar
28	568	29.0	426	17 08TNN7	08TNN7 methanosarc
29	568	29.0	410	13 073790	073790 gallus gall
30	565.5	28.9	423	13 091422	091422 gallus gall
31	564	28.8	455	4 013815	013815 homo sapien
32	555	28.3	359	11 09D6A7	09D6A7 mus musculus
33	553	28.2	459	11 09WTT1	09WTT1 cavia porce
34	550.5	28.1	424	17 08TKL5	08TKL5 methanosarc
35	542.5	27.7	448	13 09W648	09W648 fuqu rubrip
36	536.5	27.4	452	13 08VXX2	08VXX2 sphenodon p
37	535	27.3	377	5 09GPI3	09GPI3 ixodes ricl
38	510.5	26.1	380	5 08WGX0	08WGX0 thripicephal
39	496	25.3	334	12 09DHG4	09DHG4 yaba-like d
40	476.5	24.3	407	4 09UNU9	09UNU9 homo sapien
41	476	24.3	418	11 003734	003734 mus musculus
42	468.5	23.9	359	5 017365	017365 caenorhabdi
43	468.5	23.9	418	11 09D7D2	09D7D2 mus musculus
44	467.5	23.9	404	6 09N212	09N212 bos taurus
45	467	23.8	373	5 09NH65	09NH65 trichinella

ALIGNMENTS

RESULT 1
ID 09D695 PRELIMINARY; PRT; 380 AA.
AC 09D695:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 4631416M05RIK protein (Megsin).
GN SERPINB7 OR 4631416M05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SKIN;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kesukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya K., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmberg L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

RX MEDLINE-21368006; PubMed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent megalin revealed its up-regulation in
RT mesangio proliferative nephritis."
RL Kidney Int. 60:641-652(2001).
DR EMBL; AK014524; BAB29410.1; -.
DR EMBL; AF105328; AAL16768.1; -.
DR HSSP; P05121; 1A7C.
DR MGD; MGI:2151053; Serpinb7.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; Serpin; 1.
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCFB9CF4 CRC64;

Query Match 76.8%; Score 1505; DB 11; Length 380;
Best Local Similarity 73.9%; Pred. No. 1.2e-95;
Matches 281; Conservative 53; Mismatches 46; Indels 0; Gaps 0;

QY 1 MASLAANAEFCNLEFREMDDNGNGNVFSSLSLFAALALVRLGAQDLSQIDKLLHV 60
DB 1 MASLAANAEFGDLEFREMDSOGNGNVFSSLSITFALTLLRLGARGCARQIDKALHF 60
QY 61 NTASGYGNSNSOGSLQSLKRVFSIDINASHKDYLSYNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOGSLQSLKRVFSIDINASHKDYLSYNGLFAEKVYGFHKDYIECAEK 120
QY 121 LVNAKVERVDFTHLDETRNTIKKWEENTHGHKIKKVIIEGGISSSAVAVLVNAVYFKK 180
DB 121 LVNAKVERVDFTHLDETRNTIKKWEENTHGHKIKKVIIEGGISSSAVAVLVNAVYFKK 180
QY 181 MOSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKITELRYNGINMYVL 240
DB 181 WKSAFTKSDTSLCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKITELRYNGINMYVL 240
QY 241 LPEADLSIEISKLSFQNLMDWTNPRRTSKYVEVFPPQKIEKNYEMKOYLRALGLKDI 300
DB 241 LPEADLSIEISKLSFQNLMDWTNPRRTSKYVEVFPPQKIEKNYEMKOYLRALGLKDI 300
QY 301 DESKADLSGASGRIYSRMHKSLEYTEEBTEATAAGSNIVKOLPOSTLFRADHP 360
DB 301 DESKADLSGASGRIYSRMHKSLEYTEEBTEATAAGSNIVKOLPOSTLFRADHP 360
QY 361 FLFVIRKDDIIFSGRVSCP 380
DB 361 FLFVIRKDDIIFSGRVSCP 380
QY 361 FLFVIRKDDIIFSGRVSCP 380
DB 361 FLFVIRKDDIIFSGRVSCP 380

RESULT 2
Q920J5 PRELIMINARY; PRT; 380 AA.
ID 0920J5;
AC 0920J5;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, last annotation update)
DE Megsin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE-21368006; PubMed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent megalin revealed its up-regulation in
RT mesangio proliferative nephritis."
RL Kidney Int. 60:641-652(2001).
DR EMBL; AF105328; AAL16768.1; -.
DR EMBL; AK014524; BAB29410.1; -.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.

```

```

DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
KW Serpin.
SQ SEQUENCE 380 AA; 42821 MW; D8076CA8EE2C2FBC CRC64;

Query Match 76.0%; Score 1488; DB 11; Length 380;
Best Local Similarity 73.7%; Pred. No. 1.8e-94;
Matches 280; Conservative 57; Mismatches 43; Indels 0; Gaps 0;

QY 1 MASLAANAEFCNLEFREMDDNGNGNVFSSLSLFAALALVRLGAQDLSQIDKLLHV 60
DB 1 MASLAANAEFGDLEFREMDSOGNGNVFSSLSITFALTLLRLGARGCARQIDKALHF 60
QY 61 NTASGYGNSNSOGSLQSLKRVFSIDINASHKDYLSYNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOGSLQSLKRVFSIDINASHKDYLSYNGLFAEKVYGFHKDYIECAEK 120
QY 121 LVNAKVERVDFTHLDETRNTIKKWEENTHGHKIKKVIIEGGISSSAVAVLVNAVYFKK 180
DB 121 LVNAKVERVDFTHLDETRNTIKKWEENTHGHKIKKVIIEGGISSSAVAVLVNAVYFKK 180
QY 181 MOSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKITELRYNGINMYVL 240
DB 181 WKSAFTKSDTSLCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKITELRYNGINMYVL 240
QY 241 LPEADLSIEISKLSFQNLMDWTNPRRTSKYVEVFPPQKIEKNYEMKOYLRALGLKDI 300
DB 241 LPEADLSIEISKLSFQNLMDWTNPRRTSKYVEVFPPQKIEKNYEMKOYLRALGLKDI 300
QY 301 DESKADLSGASGRIYSRMHKSLEYTEEBTEATAAGSNIVKOLPOSTLFRADHP 360
DB 301 VESRADLSGASGRIYSRMHKSLEYTEEBTEATAAGSNIVKOLPOSTLFRADHP 360
QY 361 FLFVIRKDDIIFSGRVSCP 380
DB 361 FLFVIRKDDIIFSGRVSCP 380

RESULT 3
Q91212 PRELIMINARY; PRT; 388 AA.
ID 091212;
AC 091212;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, last annotation update)
DE RIKEN CDNA 2310046M08 gene.
GN 2310046M08RIK.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11;
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; BC010313; AAL10313.1; -.
DR MGD; MGI:1914207; 2310046M08RIK.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
SQ SEQUENCE 388 AA; 43544 MW; AF4BB62FB084E6FA CRC64;

Query Match 40.8%; Score 800; DB 11; Length 388;
Best Local Similarity 41.8%; Pred. No. 3.6e-47;
Matches 162; Conservative 81; Mismatches 137; Indels 8; Gaps 5;

QY 1 MASLAANAEFCNLEFREMDDNGNGNVFSSLSLFAALALVRLGAQDLSQIDKLLHV 60
DB 1 MDPITTAATPFCADVKEKLSNNVGENIFFSPLTFTYALSMILLGTRGKSABQMEKVLHY 60
QY 61 NTASGYGNSNSOGSLQSLKRVFSIDINASHKDYLSYNGLFAEKVYGFHKDYIECAEK 120
DB 61 NTASGYGNSNSOGSLQSLKRVFSIDINASHKDYLSYNGLFAEKVYGFHKDYIECAEK 120
QY 121 LVNAKVERVDFTHLDETRNTIKKWEENTHGHKIKKVIIEGGISSSAVAVLVNAVYFKK 180
DB 121 LVNAKVERVDFTHLDETRNTIKKWEENTHGHKIKKVIIEGGISSSAVAVLVNAVYFKK 180
QY 181 MOSAFTKSETINCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKITELRYNGINMYVL 240
DB 181 WKSAFTKSDTSLCHFKSPKCSKAVAMMHQERKFNLSVIEDPSMKITELRYNGINMYVL 240
QY 241 LPEADLSIEISKLSFQNLMDWTNPRRTSKYVEVFPPQKIEKNYEMKOYLRALGLKDI 300
DB 241 LPEADLSIEISKLSFQNLMDWTNPRRTSKYVEVFPPQKIEKNYEMKOYLRALGLKDI 300
QY 301 DESKADLSGASGRIYSRMHKSLEYTEEBTEATAAGSNIVKOLPOSTLFRADHP 360
DB 301 DESKADLSGASGRIYSRMHKSLEYTEEBTEATAAGSNIVKOLPOSTLFRADHP 360
QY 361 FLFVIRKDDIIFSGRVSCP 380
DB 361 FLFVIRKDDIIFSGRVSCP 380

```


OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BAB/C; TISSUE=LUNG;
 RA Kalsnerman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
 RA McKone R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
 RA Ragusea J., Bird P.I.;
 RT "Comparison of human chromosome 6p25 with murine chromosome 13 reveals
 RT a greatly expanded *Ov-serpin* gene repertoire in the mouse."
 RL Genomics 0:0-0(2002).
 DR EMBL; AF426025; AAL57487.1;
 DR InterPro: IPR002016; Peroxidase.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00435; PEROXIDASE.1; UNKNOWN.1.
 DR PROSITE: PS00284; SERPIN; UNKNOWN.1.
 SQ SEQUENCE 382 AA; 42887 MW; 35CB86ADF77C8DB CRC64;

Query Match 36.1%; Score 707; DB 11; Length 382;
 Best Local Similarity 38.8%; Pred. No. 8.7e-41;
 Matches 152; Conservative 76; Mismatches 142; Indels 22; Gaps 7;

OY 1 MASLAANAECFENLFREMDNONGNVEFSSLSFALALVRLGAODDSLSQDKLHV 60
 DB 1 MEOLSSANTLTLELFHTLKESPFGNIFFPFSSISLAWVFLGAKGSTAOLSKTFH- 59
 OY 61 NTASGYGSSNSQSGLSQSLKRVFSDINASHKDYDLSYNGLFAKVGFGHNDYIECAEK 120
 DB 60 -----FDSVEDIHSCFQSLTAEV-SKLGASH--TLKLANRLYGEKTYNLFDEPLASQK 110
 OY 121 LYDAVERVDFTNHLIEDTRRNINKVVENETHGKIRNVGEGCISSAVAVLVNAVYFPGK 180
 DB 111 MSADLAADVDFHASEDARKEINQWVKQTEGKIPELLAKGVADSMFKVLVNAVYFPGK 170
 OY 181 WOSATKSETINCHFKSPKCSGKAVAMHOKERFNLVIEDPSMKILELRYNGG-INNVY 239
 DB 171 WEQDMTRETINAPRLKKNKDKTKYKMMYQKKKFFPGYISDLCKCYVLEMPYOGGELSMT 230
 OY 240 LLPEN-----DLSEIENKLTFFONLMEWNTNPRRMTSKYVEVFPQFKIEKNEMKOYLRA 293
 DB 231 LLPEDIEDESGTKLIEQGLTGKLEHMTKHENLNIDVHVLPFRKMEESTYLSNLC 290
 OY 294 LGLKIDFESKADLSGIASGRLYISRMHKSIEVTEGTEATGATGSGNT--VEKOLP 350
 DB 291 LGVODLFSSKADLSGMSGRDLFVYSKIVHKSFDVYNEGTEAATAATGIIQVLCCKMPT 350
 OY 351 OSTFRADHPFLFVIRKDDI--ILFSGKVS 380
 DB 351 POFVTVDPHPLFTRHNPTANMIFGRVCS 382

RESULT 8
 O9D7S8 PRELIMINARY: PRT: 379 AA.
 AC O9D7S8;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE 1190005M04RIK protein.
 GN SERPINB1A OR 1190005M04RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikiado I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Stabili P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL; AK008914; BAB25964.1;
 DR HSP; P05619; IHLF.
 DR MGD; MGI:1913472; Serpinb1a.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin.1.
 DR SMART: SM00093; SERPIN.1.
 DR PROSITE: PS00284; SERPIN.1.
 KM Serpin.
 SQ SEQUENCE 379 AA; 42580 MW; 193728585077AB4B CRC64;

Query Match 35.8%; Score 700.5; DB 11; Length 379;
 Best Local Similarity 38.6%; Pred. No. 2.4e-40;
 Matches 150; Conservative 75; Mismatches 145; Indels 19; Gaps 6;

OY 1 MASLAANAECFENLFREMDNONGNVEFSSLSFALALVRLGAODDSLSQDKLHV 60
 DB 1 MEOLSSANTLTLELFHTLKESPFGNIFFPFSSISLAWVFLGAKGSTAOLSKTFH- 59
 OY 61 NTASGYGSSNSQSGLSQSLKRVFSDINASHKDYDLSYNGLFAKVGFGHNDYIECAEK 120
 DB 60 -----FDSVEDIHSPQSONAEV-SKRGASH--TLKLANRLYGEKTYNLFDEPLASQK 110
 OY 121 LYDAVERVDFTNHLIEDTRRNINKVVENETHGKIRNVGEGCISSAVAVLVNAVYFPGK 180
 DB 111 MYGADLAPDFELHASEDARKEINQWVKQTEGKIPELLAKGVADSMFKVLVNAVYFPGK 170
 OY 181 WOSATKSETINCHFKSPKCSGKAVAMHOKERFNLVIEDPSMKILELRYNGG-INNVY 239
 DB 171 WEKFMTEEDTDPAPRLSKDKTRTKYKMMYQKKKFFPGYISDLCKCYVLEMPYOGGELSMT 230
 OY 240 LLPEN-----DLSEIENKLTFFONLMEWNTNPRRMTSKYVEVFPQFKIEKNEMKOYLRA 293
 DB 231 LLPKIDIEDESGTKLIEQGLTGKLEHMTKHENLNIDVHVLPFRKMEESTYLSNLCGR 290
 OY 294 LGLKIDFESKADLSGIASGRLYISRMHKSIEVTEGTEATGATGSGNT--VEKOLP 353
 DB 291 LGVODLFSSKADLSGMSGRDLFVYSKIVHKSFDVYNEGTEAATAATGIIQVLCCKMPT 350
 OY 354 LFRADHPFLFVIRKDDI--ILFSGKVS 380
 DB 351 EFTVDPHPLFTRHNPTANMIFGRVCS 379

RESULT 9
 O9BYF7 PRELIMINARY: PRT: 369 AA.
 AC O9BYF7;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE SCCA2b.
 GN SCCA2.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Suninmi Y., Kishi F., Murakami A., Sakaguchi Y., Kato H.;
 RT "Novel forms of SCC Antigen transcripts produced by Alternative
 RT Splicing."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL -1 SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL: AB046400; BAB40773.1;
 DR HSSP: P01008; IATH.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin.
 SQ SEQUENCE 369 AA; 42287 MW; CED7A2426378DE3E CRC64;

Query Match 35.5%; Score 694.5; DB 4; Length 369;
 Best Local Similarity 39.0%; Pred. No. 6e-40;
 Matches 153; Conservative 82; Mismatches 122; Indels 35; Gaps 10;

QY 1 MASIAANAEFCNLFREMDNONGVNFSSLSLFAALALVRLGADDSLSQIDKLHV 60
 DB 1 MNSLSEANTKFMFLFOOFKRSKEN-NIFYSPISTALGVLLGADNDNAQOISKVLHF 59
 QY 61 -----NTASGYGNSNSQSG-LOSQLRVFSIDINASHKDYDLSIVNGLFAEKVGFHKDY 114
 DB 60 DQYVENTEKAATYHYDRSGVNHQFOKLTEFKSKSDNVELKANKLFEKTYQFOLEY 119
 QY 115 IECAEKLYDAKVERVDFTNLEDTNRINKWVENETNGIKVIGEGISSAVVLVNA 174
 DB 120 LDAIKKRYQSVSESTDANAPESRRKINSWVESQTEKIKINLEPPDGTIGDITLVVNA 179
 QY 175 VYFPGKQSAFPTSETINCHFKSPKSGKAVAMHQRKFLSYIEDPSKMLELRNG- 233
 DB 180 IYFGQENKFKKNT-----KEEKFPN--KVVQAVLEIPYKKG 218
 QY 234 GINMYVLLPE--NDLSEIENKLFQNLMEWTNPRMTSKYVEFFPOFKIEKNEMQYL 291
 DB 219 DLSMIVLLPMEIDGLQLEKLTAEKLEMTSLQNMETCVDLHPFKMESDLDKDTL 278
 QY 292 RALGLKIFDESKADLSGASGRLYISRMHKSIVTEBGEFATAGSNIVKOLPQ 351
 DB 279 RINGMVAIFN-GADDLISGATWSHLSVSKLHKAFAVEVBEGVAAATAVVVELSSPS 337
 QY 352 ST-LERADHPFLFYIR--KDDILFSGKVASCP 380
 DB 338 TNEFCNHPFLFYIRKNTKNSILFYGRSSP 369

RESULT 10

QY 09D105 PRELIMINARY; PRT; 387 AA.
 AC 09D105;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 18 days embryo cDNA, RIKEN full-length enriched library,
 DE clone:1110001H02, full insert sequence.
 GN SCCA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinaawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fischmann W., Gaasterland P., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
 RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-Oka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1 SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL: AK003220; BAB22650.1;
 DR HSSP: P05120; IBY7.
 DR MCD: MGI:1277952; Sccaz.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin.
 SQ SEQUENCE 387 AA; 44558 MW; AE25C02520F3B194 CRC64;

Query Match 35.0%; Score 686.5; DB 11; Length 387;
 Best Local Similarity 36.8%; Pred. No. 2.3e-39;
 Matches 143; Conservative 91; Mismatches 144; Indels 11; Gaps 6;

QY 1 MASIAANAEFCNLFREMDNONGVNFSSLSLFAALALVRLGADDSLSQIDKLHV 59
 DB 1 MIFRHADVAFVAVEMTQQL--RESKNIIFSPISMATMLMLQAGNFIQIEKVLQF 58
 QY 60 VNTASGYGNS--NSQSGLOSQLRVFSIDINASHKDYDLSIVNGLFAEKVGFHKDYIE 116
 DB 59 IETTKTKTESSECHDDENVEHQFOKLITOLKNSINDYDILKANSIYGAGFPFLQFLE 118
 QY 117 CAEKLYDAKVERVDFTNLEDTNRINKWVENETNGIKVIGEGISSAVVLVNAV 176
 DB 119 DKEYYQAKVESLDFEATSESEKINSWVESTNGIKIDLPFGSLSSSTILVNAVY 178
 QY 177 FKGQWASAFPTSETINCHFKSPKSGKAVAMHQRKFLSYIEDPSKMLELRNG-GI 235
 DB 179 FKGQWKNKFNENHTRREKFLNKNKTSKPVQMMQKNKFNESPLGDVAQVLEIPYKGD 238
 QY 236 NMYVLLPE--NDLSEIENKLFQNLMEWTNPRMTSKYVEFFPOFKIEKNEMQYLRA 293
 DB 239 SMEVLLPMEIDGLQLEEQLTDTDKLEWIKAEWNNHLELYSLPREFVEEKYDLOVPLEH 298
 QY 294 LGLKIDFDESKADLSGASGRLYISRMHKSIVTEBGEFATAGSNIVKOLPQST 353
 DB 299 MGVADADPOKADPSGSSILPGLVSKVLKSFYEVNBSGTEAAGVEVSVRAQIAE 358
 QY 354 LERADHPFLFYI--RKDDILFSGKVASCP 380
 DB 359 DFCDHPFLFYIHRMNSILFYGRICSP 387

RESULT 11

QY 008804 PRELIMINARY; PRT; 377 AA.
 AC 008804;
 DT 01-JUN-1997 (TREMBlrel. 04, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Serine proteinase inhibitor NK13.
 GN SP112.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL;
 RX MEDLINE=9736124; PubMed=9182575;
 RA Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;
 RT "A new family of 10 murine ovalbumin serpins includes two homologs of
 RT proteinase inhibitor 8 and two homologs of the granzyme B inhibitor
 RT J. Biol. Chem. 272:15434-15441(1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL;
 RA Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
 RA McKee R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
 RA Ragousis J., Bird P.I.;
 RT "Comparison of human chromosome 6p25 with murine chromosome 13 reveals
 RT a greatly expanded Ov-serpin gene repertoire in the mouse."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR EMBL: U96707; AAB57819.2; -.
 DR MGD: MGI:894688; SP112.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 KW Serpin.
 SQ SEQUENCE 377 AA; 42535 MW; 96608E8F3A41EE47A CRC64;
 Query Match 34.9%; Score 684.5; DB 11; Length 377;
 Best Local Similarity 39.4%; Pred. No. 3e-39;
 Matches 152; Conservative 76; Mismatches 143; Indels 15; Gaps 8;
 OY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
 DB 1 MDPLLEANAATFALNLKLTIGEDSSR-NVLESPISISSALVAVLGAGTTAIOQTQALSL 59
 OY 61 NTASGYGSSNSGSC-LQSOLKRVFSDINASHKDYDLSYNGLEFAEYVGFHKDYICAE 120
 DB 60 DKSGCKGRDVOH-GFOSL---TETNKGTOVLTATNRLKEKFTDILASKDCR 114
 OY 121 LYDAKVERVDFTNHLDTNRNINKWENETHGKIKVIGEGISSAVMLVNAVYFKG 180
 DB 115 FYEAEWELDFKATQDSRQHINAWAKTEDEKTELSSGVSNSNPLVNAVYFKGN 174
 OY 181 WQSAFTKSETINCFKSPKCSGKAVAMHDEKKNLSYIEDPSKMLELRYNG-INKYV 239
 DB 175 WKKQFNKEDPQEMPFNTKDVVKPVQMFQKSTFKMYVEEISTNILLPLVYGNELNMI 234
 OY 240 LLEPEN--DLSEIENKLTFOUMLMEWTNPRMTSKYVEVFPFOKTEKVENKQYLRALGLK 297
 DB 235 MLDPHEILSMVEKEITLYKFIETRLDKMEEEVEVFLPRFKLEENVDMKDVLCRLGNT 294
 OY 298 DIFDESKADLSGIASGRLYISRMHKSYLEVEEGTEATATGNSN-VEKQLPOSTLFR 356
 DB 295 DAVEEGNADSGIASKGLFLSKVIHKSIFYEVNEEGTEAATAATNANGFRCMV---YFC 351
 OY 357 ADHPFLFVIR--KDDILFSGKVS 380
 DB 352 ANHPFLFIOHSRTSGIVFCGRFSSP 377
 RESULT 12
 ID 08VH00 PRELIMINARY; PRT; 385 AA.
 AC 08VH00;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE SP312.
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B6D F1/J;
 RA Kaiserman D., Knaggs S., Scarff K.L., Gillard A., Mirza G., Cadman M.,
 RA McKee R., Denny P., Cooley J., Benarafa C., Remold-O'Donnell E.,
 RA Ragousis J., Bird P.I.;
 RT "Comparison of Human Chromosome 6p25 with Murine Chromosome 13 Reveals
 RT a greatly expanded Ov-Serpin Gene Repertoire in the Mouse."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF425084; AAL65910.1; -.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
 DR PROSITE: PS00284; SERPIN; 1.
 SQ SEQUENCE 385 AA; 43786 MW; 0820A8A26AEBB485 CRC64;
 Query Match 34.9%; Score 684.5; DB 11; Length 385;
 Best Local Similarity 39.6%; Pred. No. 3.1e-39;
 Matches 156; Conservative 75; Mismatches 140; Indels 23; Gaps 9;
 OY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
 DB 1 MDPLLEANAATFALNLKLTIGEDRSK-NVLESPISISSALVAVLGAGTTAIOQTQALSL 59
 OY 61 NTASGYGSSNSGSC-LQSOLKRVFSDINASHKDYDLSYNGLEFAEYVGFHKDYICAE 119
 DB 60 -----GKCSSESDGDVHOGFOLLSEVNKTGYSLKANRLRGEKFTDILASFKSC 113
 OY 120 KLYDAKVERVDFTNHLDTNRNINKWENETHGKIKVIGEGISSAVMLVNAVYFKG 179
 DB 114 FYEAEWELDFKATQDSRQHINAWAKTEDEKTELSSGVSNSNPLVNAVYFKGN 173
 OY 180 WQSAFTKSETINCFKSPKCSG-----KAVAMHDEKKNLSYIEDPSKMLELRYN 232
 DB 174 WKKQFNKEDPQEMPFNTKDVVKPVQMFQKSTFKMYVEEISTNILLPLVYGNELNMI 233
 OY 233 GG-INKYVLLPEN--DLSEIENKLTFOUMLMEWTNPRMTSKYVEVFPFOKTEKVENKQYLRALGLK 289
 DB 234 GNEINMIMLPDEVELSTYKEKTEHKEFTETRLDRKGGKAYVFLPWFLEBNYMKD 293
 OY 290 YLRALGLKDFDESKADLSGIASGRLYISRMHKSYLEVEEGTEATATGNSNIVEKQL 349
 DB 294 VLCKLGMTDAFEERADPSGSSKQGLFLSVIHKSIFYEVNEEGTEAATAATNANGFRCMV---YFC 351
 OY 350 POST-LPRADHPFLFVIR--KDDILFSGKVS 380
 DB 352 SRSTPCFCVNRPFIFIOHKTNETILFLGRSSP 385
 RESULT 13
 ID 092262 PRELIMINARY; PRT; 388 AA.
 AC 092262;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Squamous cell carcinoma antigen 2.
 GN SCCA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99047536; PubMed=9828132;
 RA Bartuski A.J., Kamachi Y., Schlick C., Massa H., Trask B.J.,
 RA Silverman G.A.;
 RT "A murine ortholog of the human serpin SC2 maps to chromosome 1 and
 RT inhibits chymotrypsin-like serine proteinases."
 RL Genomics 54:297-306(1998).
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.

DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin.
SQ SEQUENCE 386 AA; 44313 MW; 2A3CFA78A73AD004 CRC64;

Query Match 33.6%; Score 658; DB 11; Length 386;
Best Local Similarity 36.0%; Pred. No. 2,1e-37;
Matches 138; Conservative 92; Mismatches 141; Indels 12; Gaps 6;

```
OY 7 ANAEFCNLPREMDNNGNGVNFSSLSLPAALALVRLGAODDSLQIDKLHVNTASGY 66
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 7 ADGKFTVEMYROL--RESKDNIFFSPISMITALGMLKGAKGNTIIEKVLQCNETTEK 64
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 67 GNSNS----QSGLSQLKRVFSDINASHKDYDLSTVNGLEPAEKVYGFHKDYTECAEKLY 122
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 65 TTEKSAHCDEDDNVHEQFQKLITQLNKSNDYDILKANSIYGAKGFPLQLTFLEDIKEYY 124
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 123 DAKYRVDFTNHLEDTRRNINKVNEETHGKIKNVIGEGISSAVVAVLVNAVYFKGKMQ 182
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 125 HANVESIDFEHAAEESSEKINFSYKNETNGKINDLFPSGSLSSSTKLVLNAVYFKGRWN 184
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 183 SAFTKSETINCHFKSPKCGKAAVAMHQQERFNLSTVEDPSMKITLRLYNG-GINMYVL 241
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 185 HKFDENNTIEEMFWLNKNTSIPVPMKQRNKFMSFLEDVQAQIVEIPIYKKEISMFLR 244
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 242 PE--NDLSEIENKLTLPQNLMEWTNPRRMTSKYVEVFPPQFIENKYMKOYLRLALGKDI 299
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 245 PMEIDGKQLKQLEKQTLAAKLEWTRAEENMHLTELYLMPREFVEKEKYDLPVPLECMGVNA 304
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 300 FDESKADLGIASGRLYISRMHKSYLEVTEEGTEATAATGSNIVEKQLPOSTLFRADH 359
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 305 FDPQKADFGSGMSSNOGLVGVKVLHKSFEVNEECTEADPASGEVI-LRLAQVADFRCDH 363
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 360 PFLFVI--RKDDIILFSGKVSCP 380
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 364 PFLFFIHSKTSILFFGRISSP 386
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: July 11, 2003, 11:59:12
Job time : 167.904 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 28.3285 Seconds

(Without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-1

Perfect score: 52

Sequence: 1 SERINCHFK 9

Scoring table: BLOSUM62

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCNUS_COMB.pep.*

2: /cgn2_6/ptodata/1/paa/US05_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*

8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*

9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*

10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*

11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*

12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*

13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*

14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*

15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*

16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*

17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*

18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*

19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*

20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*

21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*

22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*

23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*

24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*

25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*

26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*

27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	52	100.0	9	US-09-140-719-1
2	52	100.0	9	US-10-091-442-1
3	52	100.0	380	US-08-091-028A-34
4	52	100.0	380	US-09-140-719-34
5	52	100.0	380	US-09-508-997A-2
6	52	100.0	380	US-09-791-537-281

7	52	100.0	380	23	US-09-936-883A-2	Sequence 2, Appl1
8 <td>52<td>100.0<th>380</th><td>24</td><td>US-10-091-442-34</td><td>Sequence 34, Appl1</td></td></td>	52 <td>100.0<th>380</th><td>24</td><td>US-10-091-442-34</td><td>Sequence 34, Appl1</td></td>	100.0 <th>380</th> <td>24</td> <td>US-10-091-442-34</td> <td>Sequence 34, Appl1</td>	380	24	US-10-091-442-34	Sequence 34, Appl1
9 <td>52<td>100.0<th>380</th><td>25</td><td>US-10-110-030-2</td><td>Sequence 2, Appl1</td></td></td>	52 <td>100.0<th>380</th><td>25</td><td>US-10-110-030-2</td><td>Sequence 2, Appl1</td></td>	100.0 <th>380</th> <td>25</td> <td>US-10-110-030-2</td> <td>Sequence 2, Appl1</td>	380	25	US-10-110-030-2	Sequence 2, Appl1
10 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-151732</td><td>Sequence 151732,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-151732</td><td>Sequence 151732,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-151732</td> <td>Sequence 151732,</td>	23	21	US-09-724-059-151732	Sequence 151732,
11 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-154252</td><td>Sequence 154252,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-154252</td><td>Sequence 154252,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-154252</td> <td>Sequence 154252,</td>	23	21	US-09-724-059-154252	Sequence 154252,
12 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-156652</td><td>Sequence 156652,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-156652</td><td>Sequence 156652,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-156652</td> <td>Sequence 156652,</td>	23	21	US-09-724-059-156652	Sequence 156652,
13 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-159172</td><td>Sequence 159172,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-159172</td><td>Sequence 159172,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-159172</td> <td>Sequence 159172,</td>	23	21	US-09-724-059-159172	Sequence 159172,
14 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-161692</td><td>Sequence 161692,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-161692</td><td>Sequence 161692,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-161692</td> <td>Sequence 161692,</td>	23	21	US-09-724-059-161692	Sequence 161692,
15 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-163972</td><td>Sequence 163972,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-163972</td><td>Sequence 163972,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-163972</td> <td>Sequence 163972,</td>	23	21	US-09-724-059-163972	Sequence 163972,
16 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-166732</td><td>Sequence 166732,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-166732</td><td>Sequence 166732,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-166732</td> <td>Sequence 166732,</td>	23	21	US-09-724-059-166732	Sequence 166732,
17 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-169012</td><td>Sequence 169012,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-169012</td><td>Sequence 169012,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-169012</td> <td>Sequence 169012,</td>	23	21	US-09-724-059-169012	Sequence 169012,
18 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-171412</td><td>Sequence 171412,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-171412</td><td>Sequence 171412,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-171412</td> <td>Sequence 171412,</td>	23	21	US-09-724-059-171412	Sequence 171412,
19 <td>39<td>75.0<td>23</td><td>21</td><td>US-09-724-059-173932</td><td>Sequence 173932,</td></td></td>	39 <td>75.0<td>23</td><td>21</td><td>US-09-724-059-173932</td><td>Sequence 173932,</td></td>	75.0 <td>23</td> <td>21</td> <td>US-09-724-059-173932</td> <td>Sequence 173932,</td>	23	21	US-09-724-059-173932	Sequence 173932,
20 <td>39<td>75.0<td>185</td><td>16</td><td>US-10-179-131-7804</td><td>Sequence 7804, Ap</td></td></td>	39 <td>75.0<td>185</td><td>16</td><td>US-10-179-131-7804</td><td>Sequence 7804, Ap</td></td>	75.0 <td>185</td> <td>16</td> <td>US-10-179-131-7804</td> <td>Sequence 7804, Ap</td>	185	16	US-10-179-131-7804	Sequence 7804, Ap
21 <td>39<td>75.0<td>185</td><td>25</td><td>US-10-179-131-7804</td><td>Sequence 14712, A</td></td></td>	39 <td>75.0<td>185</td><td>25</td><td>US-10-179-131-7804</td><td>Sequence 14712, A</td></td>	75.0 <td>185</td> <td>25</td> <td>US-10-179-131-7804</td> <td>Sequence 14712, A</td>	185	25	US-10-179-131-7804	Sequence 14712, A
22 <td>39<td>75.0<td>185</td><td>27</td><td>US-09-508-997A-4</td><td>Sequence 14712, A</td></td></td>	39 <td>75.0<td>185</td><td>27</td><td>US-09-508-997A-4</td><td>Sequence 14712, A</td></td>	75.0 <td>185</td> <td>27</td> <td>US-09-508-997A-4</td> <td>Sequence 14712, A</td>	185	27	US-09-508-997A-4	Sequence 14712, A
23 <td>39<td>75.0<td>380</td><td>19</td><td>US-09-508-997A-4</td><td>Sequence 4, Appl1</td></td></td>	39 <td>75.0<td>380</td><td>19</td><td>US-09-508-997A-4</td><td>Sequence 4, Appl1</td></td>	75.0 <td>380</td> <td>19</td> <td>US-09-508-997A-4</td> <td>Sequence 4, Appl1</td>	380	19	US-09-508-997A-4	Sequence 4, Appl1
24 <td>39<td>75.0<td>380</td><td>23</td><td>US-09-936-883A-19</td><td>Sequence 19, Appl1</td></td></td>	39 <td>75.0<td>380</td><td>23</td><td>US-09-936-883A-19</td><td>Sequence 19, Appl1</td></td>	75.0 <td>380</td> <td>23</td> <td>US-09-936-883A-19</td> <td>Sequence 19, Appl1</td>	380	23	US-09-936-883A-19	Sequence 19, Appl1
25 <td>39<td>75.0<td>380</td><td>25</td><td>US-10-110-030-4</td><td>Sequence 13518, A</td></td></td>	39 <td>75.0<td>380</td><td>25</td><td>US-10-110-030-4</td><td>Sequence 13518, A</td></td>	75.0 <td>380</td> <td>25</td> <td>US-10-110-030-4</td> <td>Sequence 13518, A</td>	380	25	US-10-110-030-4	Sequence 13518, A
26 <td>38<td>73.1<td>199</td><td>20</td><td>US-09-614-150-13518</td><td>Sequence 134773,</td></td></td>	38 <td>73.1<td>199</td><td>20</td><td>US-09-614-150-13518</td><td>Sequence 134773,</td></td>	73.1 <td>199</td> <td>20</td> <td>US-09-614-150-13518</td> <td>Sequence 134773,</td>	199	20	US-09-614-150-13518	Sequence 134773,
27 <td>38<td>73.1<td>199</td><td>21</td><td>US-09-791-537-134773</td><td>Sequence 134773,</td></td></td>	38 <td>73.1<td>199</td><td>21</td><td>US-09-791-537-134773</td><td>Sequence 134773,</td></td>	73.1 <td>199</td> <td>21</td> <td>US-09-791-537-134773</td> <td>Sequence 134773,</td>	199	21	US-09-791-537-134773	Sequence 134773,
28 <td>38<td>73.1<td>199</td><td>27</td><td>US-60-167-245-514</td><td>Sequence 514, App</td></td></td>	38 <td>73.1<td>199</td><td>27</td><td>US-60-167-245-514</td><td>Sequence 514, App</td></td>	73.1 <td>199</td> <td>27</td> <td>US-60-167-245-514</td> <td>Sequence 514, App</td>	199	27	US-60-167-245-514	Sequence 514, App
29 <td>38<td>73.1<td>199</td><td>27</td><td>US-60-191-637-13559</td><td>Sequence 13559, A</td></td></td>	38 <td>73.1<td>199</td><td>27</td><td>US-60-191-637-13559</td><td>Sequence 13559, A</td></td>	73.1 <td>199</td> <td>27</td> <td>US-60-191-637-13559</td> <td>Sequence 13559, A</td>	199	27	US-60-191-637-13559	Sequence 13559, A
30 <td>38<td>73.1<td>199</td><td>27</td><td>US-60-191-637-13559</td><td>Sequence 10688, A</td></td></td>	38 <td>73.1<td>199</td><td>27</td><td>US-60-191-637-13559</td><td>Sequence 10688, A</td></td>	73.1 <td>199</td> <td>27</td> <td>US-60-191-637-13559</td> <td>Sequence 10688, A</td>	199	27	US-60-191-637-13559	Sequence 10688, A
31 <td>38<td>73.1<td>208</td><td>20</td><td>US-09-614-150-9252</td><td>Sequence 9252, Ap</td></td></td>	38 <td>73.1<td>208</td><td>20</td><td>US-09-614-150-9252</td><td>Sequence 9252, Ap</td></td>	73.1 <td>208</td> <td>20</td> <td>US-09-614-150-9252</td> <td>Sequence 9252, Ap</td>	208	20	US-09-614-150-9252	Sequence 9252, Ap
32 <td>38<td>73.1<td>208</td><td>21</td><td>US-09-791-537-126235</td><td>Sequence 126235,</td></td></td>	38 <td>73.1<td>208</td><td>21</td><td>US-09-791-537-126235</td><td>Sequence 126235,</td></td>	73.1 <td>208</td> <td>21</td> <td>US-09-791-537-126235</td> <td>Sequence 126235,</td>	208	21	US-09-791-537-126235	Sequence 126235,
33 <td>38<td>73.1<td>208</td><td>27</td><td>US-60-167-245-386</td><td>Sequence 386, App</td></td></td>	38 <td>73.1<td>208</td><td>27</td><td>US-60-167-245-386</td><td>Sequence 386, App</td></td>	73.1 <td>208</td> <td>27</td> <td>US-60-167-245-386</td> <td>Sequence 386, App</td>	208	27	US-60-167-245-386	Sequence 386, App
34 <td>38<td>73.1<td>208</td><td>27</td><td>US-60-173-664-7495</td><td>Sequence 7495, Ap</td></td></td>	38 <td>73.1<td>208</td><td>27</td><td>US-60-173-664-7495</td><td>Sequence 7495, Ap</td></td>	73.1 <td>208</td> <td>27</td> <td>US-60-173-664-7495</td> <td>Sequence 7495, Ap</td>	208	27	US-60-173-664-7495	Sequence 7495, Ap
35 <td>38<td>73.1<td>208</td><td>27</td><td>US-60-191-637-9281</td><td>Sequence 9281, Ap</td></td></td>	38 <td>73.1<td>208</td><td>27</td><td>US-60-191-637-9281</td><td>Sequence 9281, Ap</td></td>	73.1 <td>208</td> <td>27</td> <td>US-60-191-637-9281</td> <td>Sequence 9281, Ap</td>	208	27	US-60-191-637-9281	Sequence 9281, Ap
36 <td>38<td>73.1<td>208</td><td>27</td><td>US-60-191-681-7241</td><td>Sequence 7241, Ap</td></td></td>	38 <td>73.1<td>208</td><td>27</td><td>US-60-191-681-7241</td><td>Sequence 7241, Ap</td></td>	73.1 <td>208</td> <td>27</td> <td>US-60-191-681-7241</td> <td>Sequence 7241, Ap</td>	208	27	US-60-191-681-7241	Sequence 7241, Ap
37 <td>37<td>71.2<td>9</td><td>4</td><td>US-08-091-028A-1</td><td>Sequence 1, Appl1</td></td></td>	37 <td>71.2<td>9</td><td>4</td><td>US-08-091-028A-1</td><td>Sequence 1, Appl1</td></td>	71.2 <td>9</td> <td>4</td> <td>US-08-091-028A-1</td> <td>Sequence 1, Appl1</td>	9	4	US-08-091-028A-1	Sequence 1, Appl1
38 <td>37<td>71.2<td>9</td><td>4</td><td>US-08-070-669A-5</td><td>Sequence 5, Appl1</td></td></td>	37 <td>71.2<td>9</td><td>4</td><td>US-08-070-669A-5</td><td>Sequence 5, Appl1</td></td>	71.2 <td>9</td> <td>4</td> <td>US-08-070-669A-5</td> <td>Sequence 5, Appl1</td>	9	4	US-08-070-669A-5	Sequence 5, Appl1
39 <td>37<td>71.2<td>320</td><td>5</td><td>US-08-172-528A-13</td><td>Sequence 13, Appl1</td></td></td>	37 <td>71.2<td>320</td><td>5</td><td>US-08-172-528A-13</td><td>Sequence 13, Appl1</td></td>	71.2 <td>320</td> <td>5</td> <td>US-08-172-528A-13</td> <td>Sequence 13, Appl1</td>	320	5	US-08-172-528A-13	Sequence 13, Appl1
40 <td>41</td> <td>71.2<td>320</td><td>7</td><td>US-08-384-032-4</td><td>Sequence 4, Appl1</td></td>	41	71.2 <td>320</td> <td>7</td> <td>US-08-384-032-4</td> <td>Sequence 4, Appl1</td>	320	7	US-08-384-032-4	Sequence 4, Appl1
41 <td>37<td>71.2<td>638</td><td>26</td><td>US-10-219-999-59911</td><td>Sequence 59911, A</td></td></td>	37 <td>71.2<td>638</td><td>26</td><td>US-10-219-999-59911</td><td>Sequence 59911, A</td></td>	71.2 <td>638</td> <td>26</td> <td>US-10-219-999-59911</td> <td>Sequence 59911, A</td>	638	26	US-10-219-999-59911	Sequence 59911, A
42 <td>37<td>71.2<td>1030</td><td>1</td><td>PCT-US01-229288-140</td><td>Sequence 140, App</td></td></td>	37 <td>71.2<td>1030</td><td>1</td><td>PCT-US01-229288-140</td><td>Sequence 140, App</td></td>	71.2 <td>1030</td> <td>1</td> <td>PCT-US01-229288-140</td> <td>Sequence 140, App</td>	1030	1	PCT-US01-229288-140	Sequence 140, App
43 <td>37<td>71.2<td>23</td><td>21</td><td>US-09-724-059-151372</td><td>Sequence 151372,</td></td></td>	37 <td>71.2<td>23</td><td>21</td><td>US-09-724-059-151372</td><td>Sequence 151372,</td></td>	71.2 <td>23</td> <td>21</td> <td>US-09-724-059-151372</td> <td>Sequence 151372,</td>	23	21	US-09-724-059-151372	Sequence 151372,
44 <td>36<td>69.2<td>23</td><td>21</td><td>US-09-724-059-152932</td><td>Sequence 152932,</td></td></td>	36 <td>69.2<td>23</td><td>21</td><td>US-09-724-059-152932</td><td>Sequence 152932,</td></td>	69.2 <td>23</td> <td>21</td> <td>US-09-724-059-152932</td> <td>Sequence 152932,</td>	23	21	US-09-724-059-152932	Sequence 152932,
45 <td>36<td>69.2<td>23</td><td>21</td><td>US-09-724-059-152932</td><td>Sequence 152932,</td></td></td>	36 <td>69.2<td>23</td><td>21</td><td>US-09-724-059-152932</td><td>Sequence 152932,</td></td>	69.2 <td>23</td> <td>21</td> <td>US-09-724-059-152932</td> <td>Sequence 152932,</td>	23	21	US-09-724-059-152932	Sequence 152932,

ALIGNMENTS

RESULT 1

US-09-140-719-1

Sequence 1, Application US/09140719

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

INVENTOR: IMASA, Fuyuki

APPLICANT: TSURUOKA, Nobuo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, Nobuhiko

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMACHITA, Kozo

APPLICANT: YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-1

Query Match 100.0%; Score 52; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 1 SETINCHK 9

RESULT 2
US-10-091-442-1
Sequence 1, Application US/10091442
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSURUOKA, Nobuo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-091-442-1

Query Match 100.0%; Score 52; DB 24; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 1 SETINCHK 9

RESULT 3
US-08-091-028A-34
Sequence 34, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSURUOKA, Nobuo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK

REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34

Query Match
Best Local Similarity 100.0%; Score 52; DB 4; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMACHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match
Best Local Similarity 100.0%; Score 52; DB 15; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 5
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2

Query Match
Best Local Similarity 100.0%; Score 52; DB 19; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 6
US-09-791-537-281
Sequence 281, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent In version 3.0
SEQ ID NO 281
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-281

Query Match
Best Local Similarity 100.0%; Score 52; DB 21; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 7
US-09-936-883A-2
Sequence 2, Application US/09936883A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
FILE REFERENCE: F2-10IDPICT
CURRENT APPLICATION NUMBER: US/09/936, 883A
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883A-2

Query Match 100.0%; Score 52; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 8
US-10-091-442-34
Sequence 34, Application US/10091442
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IWASA, Fuyuki
TSURUOKA, Nobuo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 52; DB 24; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 9
US-10-110-030-2
Sequence 2, Application US/10110030
GENERAL INFORMATION:
APPLICANT: Miyata, Toshio
TITLE OF INVENTION: ANIMAL MODEL FOR MESENCHYMAL PROLIFERATIVE
FILE REFERENCE: SHIM015
CURRENT APPLICATION NUMBER: US/10/110,030
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 11/285736
PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-10-110-030-2

Query Match 100.0%; Score 52; DB 25; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 10
US-09-724-059-151732
Sequence 151732, Application US/09724059
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
APPLICANT: Isalan, Mark
TITLE OF INVENTION: Nucleic Acid Binding Proteins
FILE REFERENCE: P2500USM
CURRENT APPLICATION NUMBER: US/09/724,059
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB9710809.6
PRIOR FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 1407122
SOFTWARE: MacAllister
SEQ ID NO 151732
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial sequence

FEATURE:
OTHER INFORMATION: Polypeptide sequence
US-09-724-059-151732

Query Match 75.0%: Score 39; DB 21; Length 23;
Best Local Similarity 66.7%: Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHK 9
DB 13 SETLSCHK 21

RESULT 11
US-09-724-059-154252
Sequence 154252, Application US/09724059
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
TITLE OF INVENTION: Nucleic Acid Binding Proteins
FILE REFERENCE: P2500USM
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB9710809.6
PRIOR FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 1407122
SOFTWARE: MacCallister
SEQ ID NO 154252
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide sequence
US-09-724-059-154252

Query Match 75.0%: Score 39; DB 21; Length 23;
Best Local Similarity 66.7%: Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHK 9
DB 13 SETLSCHK 21

RESULT 12
US-09-724-059-156652
Sequence 156652, Application US/09724059
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
APPLICANT: Isalan, Mark
TITLE OF INVENTION: Nucleic Acid Binding Proteins
FILE REFERENCE: P2500USM
CURRENT APPLICATION NUMBER: US/09/724,059
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB9710809.6
PRIOR FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 1407122
SOFTWARE: MacCallister
SEQ ID NO 156652
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide sequence
US-09-724-059-156652

Query Match 75.0%: Score 39; DB 21; Length 23;
Best Local Similarity 66.7%: Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHK 9

DB 13 SETLSCHK 21

RESULT 13
US-09-724-059-159172
Sequence 159172, Application US/09724059
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
APPLICANT: Isalan, Mark
TITLE OF INVENTION: Nucleic Acid Binding Proteins
FILE REFERENCE: P2500USM
CURRENT APPLICATION NUMBER: US/09/724,059
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB9710809.6
PRIOR FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 1407122
SOFTWARE: MacCallister
SEQ ID NO 159172
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide sequence
US-09-724-059-159172

Query Match 75.0%: Score 39; DB 21; Length 23;
Best Local Similarity 66.7%: Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHK 9
DB 13 SETLSCHK 21

RESULT 14
US-09-724-059-161692
Sequence 161692, Application US/09724059
GENERAL INFORMATION:
APPLICANT: Choo, Yen
APPLICANT: Klug, Aaron
APPLICANT: Isalan, Mark
TITLE OF INVENTION: Nucleic Acid Binding Proteins
FILE REFERENCE: P2500USM
CURRENT APPLICATION NUMBER: US/09/724,059
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB9710809.6
PRIOR FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 1407122
SOFTWARE: MacCallister
SEQ ID NO 161692
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide sequence
US-09-724-059-161692

Query Match 75.0%: Score 39; DB 21; Length 23;
Best Local Similarity 66.7%: Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHK 9
DB 13 SETLSCHK 21

RESULT 15
US-09-724-059-163972
Sequence 163972, Application US/09724059
GENERAL INFORMATION:
APPLICANT: Choo, Yen

APPLICANT: Klug, Aaron
APPLICANT: Isaiian, Mark
TITLE OF INVENTION: Nucleic Acid Binding Proteins
FILE REFERENCE: P2500USM
CURRENT APPLICATION NUMBER: US/09/724,059
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: GB9710809.6
PRIOR FILING DATE: 1997-05-23
NUMBER OF SEQ ID NOS: 1407122
SOFTWARE: MacAllister
SEQ ID NO: 163972
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide sequence
US-09-724-059-163972

Query Match 75.0%; Score 39; DB 21; Length 23;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
|||::|||
DB 13 SETLSCHTK 21

Search completed: July 11, 2003, 12:27:26
Job time : 31.3285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 8.45738 Seconds

(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-1

Perfect score: 52

Sequence: 1 SETINCHFX 9

Scoring table: BIOSUM62

Searched: 1229476 seqs, 263775074 residues

Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	380	10	US-09-936-883C-2
2	52	100.0	380	12	US-10-126-052A-263
3	52	100.0	380	14	US-60-443-566-3006
4	52	100.0	380	14	US-60-445-444-4692
5	52	100.0	380	14	US-60-465-241-4692
6	40	76.9	53	12	US-10-424-599-202788
7	39	75.0	380	10	US-09-936-883C-19
8	37	71.2	56	12	US-10-424-599-158716
9	37	71.2	275	12	US-10-425-114-47479
10	37	71.2	320	12	US-10-219-051B-13364
11	37	71.2	638	12	US-10-425-114-71838
12	37	71.2	1030	2	PCT-US01-29288A-140
13	37	71.2	1030	12	US-10-149-310-140
14	36	69.2	86	12	US-10-424-599-223211
15	36	69.2	239	12	US-10-425-114-58872
16	36	69.2	421	12	US-10-425-114-57566
17	36	69.2	863	2	PCT-US01-29288A-264
18	36	69.2	863	12	US-10-149-310-264
19	36	69.2	1007	12	US-10-425-114-62667

20	36	69.2	1215	12	US-10-437-963-129618	Sequence 129618, A
21	35	67.3	309	12	US-10-425-114-51109	Sequence 51109, A
22	35	67.3	522	12	US-10-282-122A-73834	Sequence 73834, A
23	35	67.3	858	12	US-10-369-294-18	Sequence 18, App1
24	35	67.3	864	12	US-10-437-963-153309	Sequence 153309, A
25	35	67.3	1173	12	US-10-369-493-5025	Sequence 5025, App
26	34	65.4	56	12	US-10-424-599-192457	Sequence 192457, A
27	34	65.4	75	12	US-10-425-114-50539	Sequence 50539, A
28	34	65.4	117	12	US-10-425-114-40164	Sequence 40164, A
29	34	65.4	120	12	US-10-416-991-111	Sequence 111, App
30	34	65.4	123	12	US-10-416-991-110	Sequence 110, App
31	34	65.4	127	12	US-10-417-884-5938	Sequence 5938, App
32	34	65.4	201	12	US-10-125-923A-608	Sequence 608, App
33	34	65.4	201	12	US-10-223-089-344	Sequence 344, App
34	34	65.4	201	12	US-10-205-892-608	Sequence 608, App
35	34	65.4	201	12	US-10-174-575-608	Sequence 608, App
36	34	65.4	201	12	US-10-187-755-608	Sequence 608, App
37	34	65.4	201	12	US-10-187-749-608	Sequence 608, App
38	34	65.4	201	12	US-10-199-672-608	Sequence 608, App
39	34	65.4	201	12	US-10-184-486-608	Sequence 608, App
40	34	65.4	201	12	US-10-305-654-344	Sequence 344, App1
41	34	65.4	215	12	US-10-360-186-54	Sequence 54, App1
42	34	65.4	292	12	US-10-425-114-39749	Sequence 39749, A
43	34	65.4	292	12	US-10-425-114-40007	Sequence 40007, A
44	34	65.4	292	12	US-10-424-599-183093	Sequence 183093, A
45	34	65.4	292	12	US-10-424-599-183093	Sequence 183093, A

ALIGNMENTS

RESULT 1
US-09-936-883C-2
; Sequence 2, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsln Protein and Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: F2-10DPIBCT
; CURRENT APPLICATION NUMBER: US/09/936.883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883C-2

Query Match 100.0%; Score 52; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFX 9
DB 188 SETINCHFX 196

RESULT 2
US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Richard
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer. Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE REFERENCE: 018501-00153005
; CURRENT APPLICATION NUMBER: US/10/126.052A

;; CURRENT FILING DATE: 2002-04-18
;; PRIOR APPLICATION NUMBER: US 60/284,770
;; PRIOR FILING DATE: 2001-04-18
;; PRIOR APPLICATION NUMBER: US 60/290,492
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: US 60/339,245
;; PRIOR FILING DATE: 2001-11-09
;; PRIOR APPLICATION NUMBER: US 60/350,666
;; PRIOR FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US 60/334,370
;; PRIOR FILING DATE: 2001-11-29
;; PRIOR APPLICATION NUMBER: US 60/372,246
;; PRIOR FILING DATE: 2002-04-12
;; NUMBER OF SEQ ID NOS: 691
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 263
;; LENGTH: 380
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-126-052A-263

Query Match
Best Local Similarity 100.0%; Score 52; DB 12; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 188 SETINCHK 196

RESULT 3
US-60-443-566-3006
;; Sequence 3006, Application US/60443566
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CLO01447
;; CURRENT APPLICATION NUMBER: US/60/443,566
;; CURRENT FILING DATE: 2003-01-30
;; NUMBER OF SEQ ID NOS: 25102
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3006
;; LENGTH: 380
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-443-566-3006

Query Match
Best Local Similarity 100.0%; Score 52; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 188 SETINCHK 196

RESULT 4
US-60-455-444-4692
;; Sequence 4692, Application US/60455444
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CLO01455
;; CURRENT APPLICATION NUMBER: US/60/455,444
;; CURRENT FILING DATE: 2003-03-18
;; NUMBER OF SEQ ID NOS: 50986
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4692
;; LENGTH: 380

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-455-444-4692

Query Match
Best Local Similarity 100.0%; Score 52; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 188 SETINCHK 196

RESULT 5
US-60-465-241-4692
;; Sequence 4692, Application US/60465241
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CLO01468
;; CURRENT APPLICATION NUMBER: US/60/465,241
;; CURRENT FILING DATE: 2003-04-23
;; NUMBER OF SEQ ID NOS: 258418
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4692
;; LENGTH: 380
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-465-241-4692

Query Match
Best Local Similarity 100.0%; Score 52; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 188 SETINCHK 196

RESULT 6
US-10-424-599-202788
;; Sequence 202788, Application US/10424599
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa Thomas J
;; APPLICANT: Kovalic David K
;; APPLICANT: Zhou Yihua
;; TITLE OF INVENTION: Soy, Nucleic Acid Molecules and Other Molecules Associated With
;; FILE REFERENCE: 38-21(53223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 285684
;; SEQ ID NO 202788
;; LENGTH: 53
;; TYPE: PRT
;; ORGANISM: Glycine max
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(53)
;; OTHER INFORMATION: unsure at all Xaa locations
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_25142C.1.pcp
US-10-424-599-202788

Query Match
Best Local Similarity 76.9%; Score 40; DB 12; Length 53;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCHK 9
: |||||

Db 35 DAINCHFK 42

RESULT 7

US-09-936-883C-19
; Sequence 19, Application US/09936883C

; GENERAL INFORMATION:

; APPLICANT: MIYATA, Toshio

; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use

; TITLE OF INVENTION: Thereof

; FILE REFERENCE: F2-10IDP1CT

; CURRENT APPLICATION NUMBER: US/09/936, 883C

; CURRENT FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: JP 1999-75305

; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: JP 1999-306623

; PRIOR FILING DATE: 1999-10-28

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 380

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 51, 94

; OTHER INFORMATION: Xaa is unknown.

US-09-936-883C-19

Query Match 75.0%; Score 39; DB 10; Length 380;

Best Local Similarity 55.6%; Pred. No. 36;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9

Db 188 SDTLSCHFR 196

RESULT 8

US-10-424-599-158716
; Sequence 158716, Application US/10424599

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424, 599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 158716

; LENGTH: 56

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MKT3847_114340C.1.pcp

US-10-424-599-158716

Query Match 71.2%; Score 37; DB 12; Length 56;

Best Local Similarity 85.7%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TINCCHK 9

Db 11 TINCCHK 17

RESULT 9

US-10-425-114-47479
; Sequence 47479, Application US/10425114

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E.

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425, 114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 47479

; LENGTH: 275

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700238237_F11.pcp

US-10-425-114-47479

Query Match 71.2%; Score 37; DB 12; Length 275;

Best Local Similarity 77.8%; Pred. No. 65;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SETINCHFK 9

Db 267 SECINCHFK 275

RESULT 10

US-10-219-051B-13964
; Sequence 13964, Application US/10219051B

; GENERAL INFORMATION:

; APPLICANT: The General Hospital Corporation doing business as Massachusetts Genera

; APPLICANT: Hospital / Bayer AG

; TITLE OF INVENTION: Nucleotide sequences involved in pain

; FILE REFERENCE: Lea 35693 Foreign Countries

; CURRENT APPLICATION NUMBER: US/10/219, 051B

; CURRENT FILING DATE: 2003-05-09

; PRIOR APPLICATION NUMBER: US 60/312,147

; PRIOR FILING DATE: 2001-08-14

; PRIOR APPLICATION NUMBER: US 60/346,382

; PRIOR FILING DATE: 2001-11-01

; PRIOR APPLICATION NUMBER: US 60/333,347

; PRIOR FILING DATE: 2001-11-26

; NUMBER OF SEQ ID NOS: 14715

; SOFTWARE: Perl script

; SEQ ID NO 13964

; LENGTH: 320

; TYPE: PRT

; ORGANISM: Rattus norvegicus

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: SWISS-Prot / CAA41937

; DATABASE ENTRY DATE: 2000-05-30

US-10-219-051B-13964

Query Match 71.2%; Score 37; DB 12; Length 320;

Best Local Similarity 55.6%; Pred. No. 73;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SETINCHFK 9

Db 163 SSTLSCHFR 171

RESULT 11

US-10-425-114-71838
; Sequence 71838, Application US/10425114

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 71838
LENGTH: 638
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: L1B3245-495-B7_FLI.pep
US-10-425-114-71838

Query Match
Best Local Similarity 71.2%; Score 37; DB 12; Length 638;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHFR 9
DB 31 SETINCHME 39

RESULT 12
PCT-US01-29288A-140
Sequence 140, Application PC/TUS0129288A
GENERAL INFORMATION:
APPLICANT: Microbia, Inc.
TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
TITLE OF INVENTION: Zinc Binuclear Cluster Proteins
FILE REFERENCE: 14184-019W01
CURRENT APPLICATION NUMBER: PCT/US01/29288A
CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 60/233,564
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 308
SOFTWARE: PatentIn version 3.1
SEQ ID NO 140
LENGTH: 1030
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
PCT-US01-29288A-140

Query Match
Best Local Similarity 71.2%; Score 37; DB 2; Length 1030;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHFR 9
DB 689 SRVLNCHFR 697

RESULT 13
US-10-149-310-140
Sequence 140, Application US/10149310
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas
APPLICANT: Madden, Kevin T.
APPLICANT: Maxon, Mary
APPLICANT: Sherman, Amir
TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
TITLE OF INVENTION: Zinc Binuclear Cluster Proteins
FILE REFERENCE: 14184-019W01
CURRENT APPLICATION NUMBER: US/10/149,310
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: PCT/US01/29288
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 60/233,564
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 308
SOFTWARE: PatentIn version 3.1

SEQ ID NO 140
LENGTH: 1030
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-149-310-140

Query Match
Best Local Similarity 71.2%; Score 37; DB 12; Length 1030;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHFR 9
DB 689 SRVLNCHFR 697

RESULT 14
US-10-424-599-223211
Sequence 223211, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 223211
LENGTH: 86
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_43590C.1.pep
US-10-424-599-223211

Query Match
Best Local Similarity 69.2%; Score 36; DB 12; Length 86;
Best Local Similarity 83.3%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 TINCHF 8
DB 56 TLNCHF 61

RESULT 15
US-10-425-114-58872
Sequence 58872, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 58872
LENGTH: 239
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700805614_FLI.pep
US-10-425-114-58872

Query Match
Best Local Similarity 69.2%; Score 36; DB 12; Length 239;
Best Local Similarity 62.3%; Pred. No. 89;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Fri Jul 11 15:10:33 2003

us-10-091-442-1.rpn

Page 5

Oy	2	ETINCHFR	9
	11	:	11:
Db	23	ETVMCHFR	30

Search completed: July 11, 2003, 12:35:04
Job time : 9.55738 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 22.0333 Seconds

(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-2

Perfect score: 40

Sequence: 1 WSAFTK 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:

1: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US089.COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US091.COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US097.COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US099.COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US102.COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US103.COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US104.COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US105.COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US106.COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US107.COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US108.COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US109.COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US110.COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US111.COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US112.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	7	US-09-140-719-2	Sequence 2, Appli
2	40	100.0	7	US-10-091-442-2	Sequence 2, Appli
3	40	100.0	380	US-08-091-028A-34	Sequence 34, Appli
4	40	100.0	380	US-09-140-719-34	Sequence 34, Appli
5	40	100.0	380	US-09-508-997A-2	Sequence 2, Appli
6	40	100.0	380	US-09-791-537-281	Sequence 281, App

7	40	100.0	380	US-09-936-883A-2	Sequence 2, Appli
8	40	100.0	380	US-10-091-442-34	Sequence 34, Appli
9	40	100.0	380	US-10-110-030-2	Sequence 2, Appli
10	36	90.0	306	PCT-US01-08631-33136	Sequence 33136, A
11	36	90.0	323	PCT-US01-08631-41856	Sequence 41856, A
12	36	90.0	368	US-09-508-997A-6	Sequence 6, Appli
13	36	90.0	368	US-09-936-883A-21	Sequence 21, Appli
14	36	90.0	368	US-10-110-030-6	Sequence 6, Appli
15	36	90.0	380	US-09-508-997A-4	Sequence 4, Appli
16	36	90.0	380	US-09-936-883A-19	Sequence 19, Appli
17	36	90.0	380	US-10-110-030-4	Sequence 4, Appli
18	35	87.5	85	US-60-160-202-2229	Sequence 2229, Ap
19	35	87.5	85	US-60-160-842-2638	Sequence 2638, Ap
20	35	87.5	95	US-60-164-763-713	Sequence 713, App
21	35	87.5	95	US-60-169-842-4466	Sequence 4466, Ap
22	35	87.5	139	PCT-US01-14827-11233	Sequence 11233, A
23	35	87.5	586	US-10-104-047-2592	Sequence 2592, Ap
24	34	85.0	54	US-60-128-476-3191	Sequence 3191, Ap
25	32	80.0	95	US-60-160-202-3165	Sequence 3165, Ap
26	32	80.0	127	US-60-169-842-4731	Sequence 4731, Ap
27	32	80.0	219	PCT-US01-08631-60256	Sequence 60256, A
28	32	80.0	331	PCT-US01-08631-48701	Sequence 48701, A
29	32	80.0	369	US-60-360-039-12166	Sequence 12166, A
30	32	80.0	404	US-10-108-260A-4601	Sequence 4601, Ap
31	32	80.0	736	US-09-791-537-70938	Sequence 70938, A
32	32	80.0	931	US-09-830-446-43	Sequence 43, Appli
33	32	80.0	940	US-09-198-452A-500	Sequence 500, App
34	32	80.0	940	US-09-438-185-468	Sequence 468, App
35	32	80.0	940	US-09-438-185A-468	Sequence 468, App
36	32	80.0	944	PCT-US01-08656-9504	Sequence 9504, Ap
37	32	80.0	944	US-60-389-987-2257	Sequence 2257, Ap
38	32	80.0	944	US-60-412-418-2257	Sequence 2257, Ap
39	32	80.0	959	US-10-109-268A-12	Sequence 12, Appli
40	32	80.0	1072	US-09-614-150-13284	Sequence 13284, A
41	32	80.0	1072	US-60-191-637-13325	Sequence 13325, A
42	32	80.0	1072	US-60-191-681-10493	Sequence 10493, A
43	32	80.0	1074	US-60-167-217-13330	Sequence 13330, A
44	32	80.0	1074	US-60-173-464-10850	Sequence 10850, A
45	31	77.5	46	PCT-US00-06012-112	Sequence 112, App

ALIGNMENTS

RESULT 1
US-09-140-719-2
; Sequence 2, Application US/09140719
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masatumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, Nobuhiko
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMACHI, Kozo
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-2

Query Match 100.0%; Score 40; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
Db 1 WOSAFTK 7

RESULT 2
US-10-091-442-2
Sequence 2, Application US/10091442
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSURUOKA, Nobuo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-091-442-2

Query Match 100.0%; Score 40; DB 24; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
Db 1 WOSAFTK 7

RESULT 3
US-08-091-028A-34
Sequence 34, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSURUOKA, Nobuo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER

REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34

Query Match 100.0%; Score 40; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 40; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 5
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2

Query Match 100.0%; Score 40; DB 19; Length 380;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 6
US-09-791-537-281
Sequence 281, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 281
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-281

Query Match 100.0%; Score 40; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 7
US-09-936-883A-2
Sequence 2, Application US/09936883A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
TITLE OF INVENTION: thereof
FILE REFERENCE: F2-101DIPCT
CURRENT APPLICATION NUMBER: US/09/936, 883A
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883A-2

Query Match
Best Local Similarity 100.0%; Score 40; DB 23; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 8
US-10-091-442-34
Sequence 34, Application US/10091442
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSUKUOKA, Nobuo
NAKAZATO, Hiroshi
MURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-Aug-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-Jun-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-Jul-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-Jul-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-Mar-1993

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match
Best Local Similarity 100.0%; Score 40; DB 24; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 9
US-10-110-030-2
Sequence 2, Application US/10110030
GENERAL INFORMATION:
APPLICANT: Miyata, Toshio
TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
FILE REFERENCE: SHIM015
CURRENT APPLICATION NUMBER: US/10/110,030
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 11/285736
PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-10-110-030-2

Query Match
Best Local Similarity 100.0%; Score 40; DB 25; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 10
PCT-US01-08631-33136
Sequence 33136, Application PC/TUS0108631
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-049
CURRENT APPLICATION NUMBER: PCT/US01/08631
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 33136
LENGTH: 306
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(306)
; OTHER INFORMATION: Xaa - x or * as defined in Table 2
PCT-US01-08631-33136

Query Match 90.0%; Score 36; DB 1; Length 306;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
Db 108 WOSVFTK 114

RESULT 11
PCT-US01-08631-41856
; Sequence 41856, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 41856
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-41856

Query Match 90.0%; Score 36; DB 1; Length 323;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
Db 108 WOSVFTK 114

RESULT 12
US-09-508-997A-6
; Sequence 6, Application US/09508997A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: Megsin Protein
; FILE REFERENCE: SHIM-004
; CURRENT APPLICATION NUMBER: US/09/508,997A
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: JP 09-275302
; PRIOR FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-508-997A-6

Query Match 90.0%; Score 36; DB 19; Length 368;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
Db 169 WKSFTK 175

RESULT 13
US-09-936-883A-21
; Sequence 21, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DP1CT
; CURRENT APPLICATION NUMBER: US/09/936,883A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-936-883A-21

Query Match 90.0%; Score 36; DB 23; Length 368;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
Db 169 WKSFTK 175

RESULT 14
US-10-110-030-6
; Sequence 6, Application US/10110030
; GENERAL INFORMATION:
; APPLICANT: Miyata, Toshio
; TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
; FILE REFERENCE: SHIM015
; CURRENT APPLICATION NUMBER: US/10/110,030
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 11/285736
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-110-030-6

Query Match 90.0%; Score 36; DB 25; Length 368;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
Db 169 WKSFTK 175

RESULT 15
US-09-508-997A-4
; Sequence 4, Application US/09508997A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: Megsin Protein
; FILE REFERENCE: SHIM-004
; CURRENT APPLICATION NUMBER: US/09/508,997A
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: JP 09-275302
; PRIOR FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentln Ver. 2.0

Fri Jul 11 15:10:39 2003

us-10-091-442-2.ram

Page 6

```

: SEQ ID NO 4
: LENGTH: 380
: TYPE: prt
: ORGANISM: Rattus rattus
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: 51..94
: OTHER INFORMATION: xaa is unknown.
US-03-508-97A-4

```

Query Match	90.0%	Score 36;	DB 19;	Length 380;
Best Local Similarity	85.7%;	Pred. NO. 2.3e+02;		
Matches	6;	Conservative	1;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	1	WQSAFTK	7
		:	
Db	181	WKSAAFTK	187

Search completed: July 11, 2003, 12:27:28
Job time : 24.0333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 6.57796 Seconds
(Without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-2

Perfect score: 40
Sequence: 1 MOSAFK 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues

Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New: *
1: /cgn2_6/ptodata/1/paa/PCY_NEW_COMB.pep: *
2: /cgn2_6/ptodata/1/paa/PCY_NEW_COMB.pep: *
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep: *
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep: *
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep: *
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep: *
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep: *
8: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep: *
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep: *
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep: *
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep: *
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep: *
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep: *
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	380	10 US-09-936-883C-2	Sequence 2, Appli
2	40	100.0	380	12 US-10-126-052A-263	Sequence 263, App
3	40	100.0	380	14 US-60-443-566-3006	Sequence 3006, App
4	40	100.0	380	14 US-60-455-444-4692	Sequence 4692, App
5	40	100.0	380	14 US-60-465-241-4692	Sequence 4692, App
6	36	90.0	368	10 US-09-936-883C-21	Sequence 21, Appl
7	36	90.0	380	10 US-09-936-883C-19	Sequence 19, Appl
8	34	85.0	1717	10 US-09-958-359-20	Sequence 20, Appl
9	33	82.5	60	12 US-10-424-599-183524	Sequence 183524,
10	32	80.0	72	12 US-10-424-599-172361	Sequence 172361,
11	32	80.0	369	12 US-10-369-493-12166	Sequence 12166, A
12	32	80.0	692	14 US-60-452-680-15528	Sequence 15528, A
13	32	80.0	938	12 US-10-282-122A-77952	Sequence 77952, A
14	32	80.0	938	12 US-10-312-273-5	Sequence 5, Appli
15	32	80.0	944	12 US-10-289-762-500	Sequence 500, App
16	32	80.0	944	10 US-09-949-016-6650	Sequence 6650, App
17	32	80.0	944	11 US-10-273-573-9504	Sequence 9504, App
18	32	80.0	944	12 US-10-218-140-6230	Sequence 6230, App
19	32	80.0	959	10 US-09-949-016-6904	Sequence 6904, App

20	32	80.0	959	12	US-10-219-051B-11182	Sequence 11182, A
21	32	80.0 <td>981</td> <td>10<td>US-09-949-016-11647</td><td>Sequence 11647, A</td></td>	981	10 <td>US-09-949-016-11647</td> <td>Sequence 11647, A</td>	US-09-949-016-11647	Sequence 11647, A
22	32	80.0 <td>981</td> <td>10<td>US-09-949-016-11648</td><td>Sequence 11648, A</td></td>	981	10 <td>US-09-949-016-11648</td> <td>Sequence 11648, A</td>	US-09-949-016-11648	Sequence 11648, A
23	32	80.0 <td>981</td> <td>12<td>US-10-087-192-624</td><td>Sequence 624, App</td></td>	981	12 <td>US-10-087-192-624</td> <td>Sequence 624, App</td>	US-10-087-192-624	Sequence 624, App
24	31	77.5 <td>93</td> <td>12<td>US-10-424-599-145438</td><td>Sequence 145438,</td></td>	93	12 <td>US-10-424-599-145438</td> <td>Sequence 145438,</td>	US-10-424-599-145438	Sequence 145438,
25	31	77.5 <td>94</td> <td>12<td>US-10-424-599-195376</td><td>Sequence 195376,</td></td>	94	12 <td>US-10-424-599-195376</td> <td>Sequence 195376,</td>	US-10-424-599-195376	Sequence 195376,
26	31	77.5 <td>127</td> <td>12<td>US-10-424-599-257209</td><td>Sequence 257209,</td></td>	127	12 <td>US-10-424-599-257209</td> <td>Sequence 257209,</td>	US-10-424-599-257209	Sequence 257209,
27	31	77.5 <td>221</td> <td>12<td>US-10-219-382-1236</td><td>Sequence 1236, App</td></td>	221	12 <td>US-10-219-382-1236</td> <td>Sequence 1236, App</td>	US-10-219-382-1236	Sequence 1236, App
28	31	77.5 <td>221</td> <td>12<td>US-10-219-382A-1236</td><td>Sequence 1236, App</td></td>	221	12 <td>US-10-219-382A-1236</td> <td>Sequence 1236, App</td>	US-10-219-382A-1236	Sequence 1236, App
29	31	77.5 <td>242</td> <td>14<td>US-60-452-680-21389</td><td>Sequence 21389, A</td></td>	242	14 <td>US-60-452-680-21389</td> <td>Sequence 21389, A</td>	US-60-452-680-21389	Sequence 21389, A
30	31	77.5 <td>242</td> <td>14<td>US-60-470-166-1193</td><td>Sequence 1193, App</td></td>	242	14 <td>US-60-470-166-1193</td> <td>Sequence 1193, App</td>	US-60-470-166-1193	Sequence 1193, App
31	31	77.5 <td>260</td> <td>12<td>US-10-437-963-125687</td><td>Sequence 125687,</td></td>	260	12 <td>US-10-437-963-125687</td> <td>Sequence 125687,</td>	US-10-437-963-125687	Sequence 125687,
32	31	77.5 <td>264</td> <td>12<td>US-10-289-762-486</td><td>Sequence 486, App</td></td>	264	12 <td>US-10-289-762-486</td> <td>Sequence 486, App</td>	US-10-289-762-486	Sequence 486, App
33	31	77.5 <td>282</td> <td>12<td>US-10-437-963-125689</td><td>Sequence 125689,</td></td>	282	12 <td>US-10-437-963-125689</td> <td>Sequence 125689,</td>	US-10-437-963-125689	Sequence 125689,
34	31	77.5 <td>514</td> <td>31</td> <td>US-10-369-493-21618</td> <td>Sequence 21618, A</td>	514	31	US-10-369-493-21618	Sequence 21618, A
35	31	77.5 <td>733</td> <td>12<td>US-10-289-762-791</td><td>Sequence 791, App</td></td>	733	12 <td>US-10-289-762-791</td> <td>Sequence 791, App</td>	US-10-289-762-791	Sequence 791, App
36	31	77.5 <td>978</td> <td>10<td>US-09-841-260-65</td><td>Sequence 65, Appl</td></td>	978	10 <td>US-09-841-260-65</td> <td>Sequence 65, Appl</td>	US-09-841-260-65	Sequence 65, Appl
37	31	77.5 <td>978</td> <td>12<td>US-10-312-273-35</td><td>Sequence 35, Appl</td></td>	978	12 <td>US-10-312-273-35</td> <td>Sequence 35, Appl</td>	US-10-312-273-35	Sequence 35, Appl
38	31	77.5 <td>1108</td> <td>12<td>US-10-282-122A-60476</td><td>Sequence 60476, A</td></td>	1108	12 <td>US-10-282-122A-60476</td> <td>Sequence 60476, A</td>	US-10-282-122A-60476	Sequence 60476, A
39	30	75.0 <td>33</td> <td>12<td>US-10-424-599-223477</td><td>Sequence 223477,</td></td>	33	12 <td>US-10-424-599-223477</td> <td>Sequence 223477,</td>	US-10-424-599-223477	Sequence 223477,
40	30	75.0 <td>75</td> <td>12<td>US-10-424-599-204603</td><td>Sequence 204603,</td></td>	75	12 <td>US-10-424-599-204603</td> <td>Sequence 204603,</td>	US-10-424-599-204603	Sequence 204603,
41	30	75.0 <td>80</td> <td>12<td>US-10-424-599-187032</td><td>Sequence 187032,</td></td>	80	12 <td>US-10-424-599-187032</td> <td>Sequence 187032,</td>	US-10-424-599-187032	Sequence 187032,
42	30	75.0 <td>88</td> <td>10<td>US-09-864-408A-2936</td><td>Sequence 2936, App</td></td>	88	10 <td>US-09-864-408A-2936</td> <td>Sequence 2936, App</td>	US-09-864-408A-2936	Sequence 2936, App
43	30	75.0 <td>91</td> <td>10<td>US-09-864-408A-4728</td><td>Sequence 4728, App</td></td>	91	10 <td>US-09-864-408A-4728</td> <td>Sequence 4728, App</td>	US-09-864-408A-4728	Sequence 4728, App
44	30	75.0 <td>130</td> <td>2<td>PCT-US02-32727-15086</td><td>Sequence 15086, A</td></td>	130	2 <td>PCT-US02-32727-15086</td> <td>Sequence 15086, A</td>	PCT-US02-32727-15086	Sequence 15086, A
45	30	75.0 <td>130</td> <td>10<td>US-09-978-825-15086</td><td>Sequence 15086, A</td></td>	130	10 <td>US-09-978-825-15086</td> <td>Sequence 15086, A</td>	US-09-978-825-15086	Sequence 15086, A

ALIGNMENTS

RESULT 1
US-09-936-883C-2
Sequence 2, Application US/09936883C
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
FILE OF INVENTION: Theoret
FILE REFERENCE: F2-101DIPACT
CURRENT APPLICATION NUMBER: US/09/936,883C
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883C-2

Query Match 100.0%; Score 40; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOSAFK 7
DB 181 MOSAFK 187

RESULT 2
US-10-126-052A-263
Sequence 263, Application US/10126052A
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Murray, Richard
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer. Compositions and
FILE REFERENCE: 018501-001530US
CURRENT APPLICATION NUMBER: US/10/126,052A

;; CURRENT FILING DATE: 2002-04-18
;; PRIOR APPLICATION NUMBER: US 60/284,770
;; PRIOR FILING DATE: 2001-04-18
;; PRIOR APPLICATION NUMBER: US 60/290,492
;; PRIOR FILING DATE: 2001-05-10
;; PRIOR APPLICATION NUMBER: US 60/339,245
;; PRIOR FILING DATE: 2001-11-09
;; PRIOR APPLICATION NUMBER: US 60/350,666
;; PRIOR FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US 60/334,370
;; PRIOR FILING DATE: 2001-11-29
;; PRIOR APPLICATION NUMBER: US 60/372,246
;; PRIOR FILING DATE: 2002-04-12
;; NUMBER OF SEQ ID NOS: 691
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 263
;; LENGTH: 380
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-126-052A-263

Query Match
Best Local Similarity 100.0%; Score 40; DB 12; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 3
US-60-443-566-3006
;; Sequence 3006, Application US/60443566
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CLO01447
;; CURRENT APPLICATION NUMBER: US/60/443,566
;; CURRENT FILING DATE: 2003-01-30
;; NUMBER OF SEQ ID NOS: 25102
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3006
;; LENGTH: 380
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-443-566-3006

Query Match
Best Local Similarity 100.0%; Score 40; DB 14; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 4
US-60-455-444-4692
;; Sequence 4692, Application US/60455444
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CLO01455
;; CURRENT APPLICATION NUMBER: US/60/455,444
;; CURRENT FILING DATE: 2003-03-18
;; NUMBER OF SEQ ID NOS: 50986
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4692
;; LENGTH: 380

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-455-444-4692

Query Match
Best Local Similarity 100.0%; Score 40; DB 14; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 5
US-60-465-241-4692
;; Sequence 4692, Application US/60465241
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CLO01468
;; CURRENT APPLICATION NUMBER: US/60/465,241
;; CURRENT FILING DATE: 2003-04-23
;; NUMBER OF SEQ ID NOS: 258418
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4692
;; LENGTH: 380
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-465-241-4692

Query Match
Best Local Similarity 100.0%; Score 40; DB 14; Length 380;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 181 WSAFTK 187

RESULT 6
US-09-936-883C-21
;; Sequence 21, Application US/09936883C
;; GENERAL INFORMATION:
;; APPLICANT: MIYATA, Toshio
;; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
;; FILE REFERENCE: F2-101DIPCT
;; CURRENT APPLICATION NUMBER: US/09/936,883C
;; CURRENT FILING DATE: 2001-12-21
;; PRIOR APPLICATION NUMBER: JP 1999-75305
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: JP 1999-306623
;; PRIOR FILING DATE: 1999-10-28
;; NUMBER OF SEQ ID NOS: 21
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 21
;; LENGTH: 368
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-936-883C-21

Query Match
Best Local Similarity 90.0%; Score 36; DB 10; Length 368;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
Db 169 WSAFTK 175

RESULT 7

RESULT 9
 US-10-424-599-183524
 : Sequence 183524, Application US/10424599
 : GENERAL INFORMATION:
 : APPLICANT: La Rosa Thomas J
 : APPLICANT: Kovalic David K
 : APPLICANT: Zhou Yihua
 : APPLICANT: Cao Yongwei
 : TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 : TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 : FILE REFERENCE: 38-21(53223)B

```

: RESULT 11
: US-10-369-493-12166
: Sequence 12166, Application US/10369493
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xianfeng
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)B
: CURRENT APPLICATION NUMBER: US/10/369,493
: PRIORITY FILING DATE: 2003-02-28
: PRIOR APPLICATION NUMBER: US 60/360,039
: PRIORITY FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 12166

```

```
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-12166

Query Match
Best Local Similarity 80.0%; Score 32; DB 12; Length 369;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
Db 315 WOSAFTK 321

RESULT 12
US-60-452-680-15528
; Sequence 15528, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001450
; CURRENT APPLICATION NUMBER: US/60/452,680
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15528
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-452-680-15528

Query Match
Best Local Similarity 80.0%; Score 32; DB 14; Length 692;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
Db 460 WSSSFTK 466

RESULT 13
US-10-282-122A-77952
; Sequence 77952, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77952
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-77952

Query Match
Best Local Similarity 80.0%; Score 32; DB 12; Length 837;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFT 6
Db 220 WOSAFT 225

RESULT 14
US-10-312-273-5
; Sequence 5, Application US/10312273
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035MO
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: Seqwin99, version 1.02
; SEQ ID NO 5
; LENGTH: 938
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-5

Query Match
Best Local Similarity 80.0%; Score 32; DB 12; Length 938;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFT 6
Db 642 WOSAFT 647

RESULT 15
US-10-289-762-500
; Sequence 500, Application US/10289762
; GENERAL INFORMATION:
```

; APPLICANT: Griffiths, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/10/289, 762
 ; CURRENT FILING DATE: 2003-03-27
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 500
 ; LENGTH: 940
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-289-762-500

Query Match 80.0% Score 32; DB 12; Length 940;
 Best Local Similarity 83.3%; Pred. NO. 5.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAFT 6
 ||||:|
 Db 644 WQSAFT 649

Search completed: July 11, 2003, 12:35:05
 Job time : 7.67796 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 59.8046 Seconds
(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-3
Perfect score: 100
Sequence: 1 VERVDFTNHLDTRRNINK 19

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	19	US-09-140-719-3	Sequence 3, Appli
2	100	100.0	19	US-10-091-442-3	Sequence 3, Appli
3	100	100.0	41	US-60-196-718-3839	Sequence 3839, Ap
4	100	100.0	380	US-08-091-028A-34	Sequence 34, Appli
5	100	100.0	380	US-09-140-719-34	Sequence 34, Appli
6	100	100.0	380	US-09-508-997A-2	Sequence 2, Appli

7	100	100.0	380	21	US-09-791-537-281	Sequence 281, App
8	100	100.0	380	23	US-09-936-883A-2	Sequence 2, Appli
9	100	100.0	380	24	US-10-091-442-34	Sequence 34, Appli
10	100	100.0	380	25	US-10-110-030-2	Sequence 2, Appli
11	82	82.0	19	4	US-08-091-028-3	Sequence 3, Appli
12	82	82.0	19	4	US-08-091-028A-3	Sequence 3, Appli
13	71	71.0	368	19	US-09-508-997A-6	Sequence 6, Appli
14	71	71.0	368	23	US-09-936-883A-21	Sequence 21, Appli
15	71	71.0	368	25	US-10-110-030-6	Sequence 6, Appli
16	68	68.0	380	13	US-09-508-997A-4	Sequence 4, Appli
17	68	68.0	380	19	US-09-936-883A-19	Sequence 19, Appli
18	68	68.0	380	25	US-10-110-030-4	Sequence 4, Appli
19	51.5	51.5	509	21	US-09-791-537-14188	Sequence 14188, A
20	51.5	51.5	509	21	US-09-791-537-88278	Sequence 88278, A
21	50	50.0	338	27	US-60-389-987-1574	Sequence 1574, Ap
22	50	50.0	338	27	US-60-412-418-1574	Sequence 1574, Ap
23	50	50.0	390	1	PCT-US02-07215-12	Sequence 12, Appli
24	50	50.0	390	4	US-08-099-259-2	Sequence 2, Appli
25	50	50.0	390	9	US-08-568-147A-2	Sequence 2, Appli
26	50	50.0	390	11	US-08-731-566-2	Sequence 2, Appli
27	50	50.0	390	20	US-09-653-464B-9	Sequence 9, Appli
28	50	50.0	390	21	US-09-791-537-33386	Sequence 33386, A
29	50	50.0	390	21	US-09-791-537-76123	Sequence 76123, A
30	50	50.0	390	21	US-09-791-537-121576	Sequence 121576, A
31	50	50.0	390	22	US-09-823-187-29	Sequence 29, Appli
32	50	50.0	390	22	US-09-823-187-30	Sequence 30, Appli
33	50	50.0	390	24	US-10-037-417-65	Sequence 65, Appli
34	50	50.0	390	24	US-10-094-944-12	Sequence 12, Appli
35	50	50.0	413	21	US-09-760-492-106	Sequence 106, App
36	50	50.0	413	25	US-09-162-749-106	Sequence 106, App
37	49	49.0	226	16	US-09-252-991A-26823	Sequence 26823, A
38	48.5	48.5	85	21	US-09-758-438-934	Sequence 934, App
39	48.5	48.5	85	26	US-10-212-760-934	Sequence 934, App
40	48.5	48.5	163	21	US-09-758-438-304	Sequence 304, App
41	48.5	48.5	163	26	US-10-212-215-304	Sequence 304, App
42	48	48.0	480	26	US-10-219-999-52273	Sequence 52273, A
43	48	48.0	480	27	US-60-324-109-28371	Sequence 28371, A
44	48	48.0	688	26	US-10-219-999-32900	Sequence 32900, A
45	48	48.0	688	27	US-60-324-109-18095	Sequence 18095, A

ALIGNMENTS

RESULT 1
US-09-140-719-3
; Sequence 3, Application US/09140719
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, Nobuhiko
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-3

Query Match 100.0%; Score 100; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNK 19
Db 1 VERVDFTNHLDETRRNK 19

RESULT 2

US-10-091-442-3

Sequence 3, Application US/10091442

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

TSURUOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhito

KURIHARA, Tatsuya

YAMAICHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-091-442-3

Query Match 100.0%; Score 100; DB 24; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNK 19
Db 1 VERVDFTNHLDETRRNK 19

RESULT 3

US-60-196-718-3839

Sequence 3839, Application US/60196718

GENERAL INFORMATION:

APPLICANT: Bonazzi, Vivien

TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

FILE REFERENCE: CL000456

CURRENT APPLICATION NUMBER: US/60/196,718

CURRENT FILING DATE: 2000-04-13

NUMBER OF SEQ ID NOS: 7494

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3839

LENGTH: 41

TYPE: PRT

ORGANISM: HUMAN

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(41)

OTHER INFORMATION: Xaa = Any Amino Acid

US-60-196-718-3839

Query Match 100.0%; Score 100; DB 27; Length 41;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNK 19
Db 15 VERVDFTNHLDETRRNK 33

RESULT 4

US-08-091-028A-34

Sequence 34, Application US/08091028A

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

APPLICANT: TSURUOKA, Nobuo

APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34

Query Match 100.0%; Score 100; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFNHLNLEDTNRNINK 19
DB 126 VERVDFNHLNLEDTNRNINK 144

RESULT 5
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masaaki
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAUCHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria

STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 100; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFNHLNLEDTNRNINK 19
DB 126 VERVDFNHLNLEDTNRNINK 144

RESULT 6
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: US/09/508,997A
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2

Query Match 100.0%; Score 100; DB 19; Length 380;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFNHLNLEDTNRNINK 19
DB 126 VERVDFNHLNLEDTNRNINK 144

Db 126 VERVDFTNHLDETRRNINK 144

RESULT 7

US-09-791-537-281
; Sequence 281, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 281
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-281

Query Match 100.0%; Score 100; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 9, 1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNINK 19
Db 126 VERVDFTNHLDETRRNINK 144

RESULT 8

US-09-936-883A-2
; Sequence 2, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Magsin Protein and Use
; FILE REFERENCE: P2-101DPICT
; CURRENT APPLICATION NUMBER: US/09/936,883A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883A-2

Query Match 100.0%; Score 100; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 9, 1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNINK 19
Db 126 VERVDFTNHLDETRRNINK 144

RESULT 9

US-10-091-442-34
; Sequence 34, Application US/10091442
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, Nobuhiko

KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS: 34
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 100; DB 24; Length 380;
Best Local Similarity 100.0%; Pred. No. 9, 1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNINK 19
Db 126 VERVDFTNHLDETRRNINK 144

RESULT 10

US-10-110-030-2
; Sequence 2, Application US/10110030
; GENERAL INFORMATION:
; APPLICANT: Miyata, Toshio
; TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
; FILE REFERENCE: SHIM015
; CURRENT APPLICATION NUMBER: US/10/110,030
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 11/285736
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 380

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-110-030-2

Query Match 100.0%; Score 100; DB 25; Length 380;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VERVDFTNHLDTNRNINK 19
Db 126 VERVDFTNHLDTNRNINK 144

RESULT 11
US-08-091-028-3
; Sequence 3, Application US/08091028
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, Nobuhiko
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/091.028
; FILING DATE: 14-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-091-028-3

Query Match 82.0%; Score 82; DB 4; Length 19;
Best Local Similarity 84.2%; Pred. No. 2.3e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VERVDFTNHLDTNRNINK 19
Db 1 VEXVDFTNHLDTXXNINK 19

RESULT 12
US-08-091-028A-3
; Sequence 3, Application US/08091028A
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, Nobuhiko
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/091.028A
; FILING DATE: 14-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: REA, TERESA STANER
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-6620
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-091-028A-3

Query Match 82.0%; Score 82; DB 4; Length 19;
Best Local Similarity 84.2%; Pred. No. 2.3e-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VERVDFTNHLDTNRNINK 19
Db 1 VEXVDFTNHLDTXXNINK 19

RESULT 13
US-09-508-997A-6
; Sequence 6, Application US/09508997A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: Megsin Protein
; FILE REFERENCE: SHIM-004
; CURRENT APPLICATION NUMBER: US/09/508.997A
; CURRENT FILING DATE: 2000-06-27

;; PRIOR APPLICATION NUMBER: JP 09-275302
;; PRIOR FILING DATE: 1997-09-22
;; NUMBER OF SEQ ID NOS: 44
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 6
;; LENGTH: 368
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-508-997A-6

Query Match 71.0%; Score 71; DB 19; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0065;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VERVDFTNHLEDTRRNK 19
|||||:||||
DB 114 VERVDFTNVDTRFRK 132

RESULT 14
US-09-936-883A-21
; Sequence 21, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: P2-10IDPICT
; CURRENT APPLICATION NUMBER: US/09/936, 883A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-936-883A-21

Query Match 71.0%; Score 71; DB 23; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0065;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VERVDFTNHLEDTRRNK 19
|||||:||||
DB 114 VERVDFTNVDTRFRK 132

RESULT 15
US-10-110-030-6
; Sequence 6, Application US/10110030
; GENERAL INFORMATION:
; APPLICANT: Miyata, Toshio
; TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
; FILE REFERENCE: SHIM015
; CURRENT APPLICATION NUMBER: US/10/110, 030
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 11/285736
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-110-030-6

Query Match 71.0%; Score 71; DB 25; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0065;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 VERVDFTNHLEDTRRNK 19
|||||:||||
DB 114 VERVDFTNVDTRFRK 132

Search completed: July 11, 2003, 12:27:30
Job time : 61.8046 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 : Search time 17.8545 Seconds

(Without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-3

Perfect score: 100

Sequence: 1 VERVDFTHLEDTFRNRINK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues

Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US12_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US13_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US14_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US15_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US16_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US17_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	380	US-09-936-883C-2	Sequence 2, Appl1
2	100	100.0	380	US-10-126-052A-263	Sequence 263, App
3	100	100.0	380	US-60-443-566-3006	Sequence 3006, App
4	100	100.0	380	US-60-455-444-4692	Sequence 4692, App
5	100	100.0	380	US-60-465-241-4692	Sequence 4692, App
6	71	71.0	368	US-09-936-883C-21	Sequence 21, Appl1
7	68	68.0	380	US-09-936-883C-19	Sequence 19, Appl1
8	50	50.0	328	US-60-443-566-4373	Sequence 4373, App
9	50	50.0	328	US-60-455-444-7455	Sequence 7455, App
10	50	50.0	328	US-60-465-241-7455	Sequence 7455, App
11	50	50.0	390	PCT-US03-17409-175	Sequence 175, Appl1
12	50	50.0	390	PCT-US02-072158-12	Sequence 12, Appl1
13	50	50.0	390	US-60-448-285-5	Sequence 5, Appl1
14	50	50.0	410	US-09-724-676-57391	Sequence 57391, A
15	50	50.0	410	US-09-724-676A-57391	Sequence 57391, A
16	49	49.0	226	US-10-366-683-26823	Sequence 26823, A
17	49	49.0	226	US-10-419-128-26823	Sequence 26823, A
18	49	49.0	1169	US-10-282-122A-47218	Sequence 47218, A
19	48.5	48.5	235	US-09-949-016-11482	Sequence 11482, A

20	48	48.0	481	US-10-425-114-68385	Sequence 68385, A
21	48	48.0	689	US-10-425-114-45472	Sequence 45472, A
22	47	47.0	503	PCT-US02-20350-4	Sequence 4, Appl1
23	47	47.0	503	US-10-219-051B-14336	Sequence 14336, A
24	47	47.0	504	PCT-US02-20350-6	Sequence 6, Appl1
25	47	47.0	504	US-10-219-051B-14338	Sequence 14338, A
26	47	47.0	504	US-60-452-680-24158	Sequence 24158, A
27	47	47.0	504	US-60-453-135-15050	Sequence 15050, A
28	47	47.0	504	US-60-453-050-15050	Sequence 15050, A
29	47	47.0	504	US-60-466-412-15050	Sequence 15050, A
30	47	47.0	506	US-10-373-877-18	Sequence 18, Appl1
31	47	47.0	507	US-09-949-016-110802	Sequence 10802, A
32	47	47.0	1043	US-09-949-016-11650	Sequence 11650, A
33	46	46.0	215	US-60-443-566-4034	Sequence 4034, App
34	46	46.0	215	US-60-455-444-5240	Sequence 5240, App
35	46	46.0	215	US-60-465-241-5240	Sequence 5240, App
36	46	46.0	258	US-60-443-566-4033	Sequence 4033, App
37	46	46.0	258	US-60-455-444-5239	Sequence 5239, App
38	46	46.0	258	US-60-465-241-5239	Sequence 5239, App
39	46	46.0	375	PCT-US02-19592-93	Sequence 93, Appl1
40	46	46.0	375	US-10-126-052A-36	Sequence 36, Appl1
41	46	46.0	375	US-10-376-564-3	Sequence 3, Appl1
42	46	46.0	375	US-10-376-564-4	Sequence 4, Appl1
43	46	46.0	375	US-10-376-564-5	Sequence 5, Appl1
44	46	46.0	375	US-60-438-735-166	Sequence 166, App
45	46	46.0	375	US-60-449-629-149	Sequence 149, App

ALIGNMENTS

RESULT 1
US-09-936-883C-2
Sequence 2, Application US/09936883C
GENERAL INFORMATION:
APPLICANT: MIVATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsln Protein and Use
FILE REFERENCE: F2-101DIPACT
CURRENT APPLICATION NUMBER: US/09/936, 883C
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883C-2

Query Match 100.0%; Score 100; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTHLEDTFRNRINK 19
|||||
DB 126 VERVDFTHLEDTFRNRINK 144

RESULT 2
US-10-126-052A-263
Sequence 263, Application US/10126052A
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Murray, Richard
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
FILE REFERENCE: 018501-001530US
CURRENT APPLICATION NUMBER: US/10/126, 052A

```

; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/339,245
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 263
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-052A-263
```

```
Query Match          100.0%; Score 100; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.2e+08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VERVDFTNHLDETRRINK 19
        |||
Db      126 VERVDFTNHLDETRRINK 144
```

```

RESULT 3
US-60-443-566-3006
; Sequence 3006, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3006
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-3006
```

```
Query Match          100.0%; Score 100; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.2e+08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VERVDFTNHLDETRRINK 19
        |||
Db      126 VERVDFTNHLDETRRINK 144
```

```

RESULT 4
US-60-455-444-4692
; Sequence 4692, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001453
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-4692
```

```
Query Match          100.0%; Score 100; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.2e+08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VERVDFTNHLDETRRINK 19
        |||
Db      126 VERVDFTNHLDETRRINK 144
```

```

RESULT 5
US-60-465-241-4692
; Sequence 4692, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-4692
```

```
Query Match          100.0%; Score 100; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.2e+08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VERVDFTNHLDETRRINK 19
        |||
Db      126 VERVDFTNHLDETRRINK 144
```

```

RESULT 6
US-09-936-883C-21
; Sequence 21, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: P2-101D1PCT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-936-883C-21
```

```
Query Match          71.0%; Score 71; DB 10; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0034;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 VERVDFTNHLDETRRINK 19
        |||
Db      114 VERVDFTNVDOTRKRINK 132
```

```
RESULT 7
```

```
US-09-936-883C-19
; Sequence 19, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DPI-PCT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 51, 94
; OTHER INFORMATION: Xaa is unknown.
US-09-936-883C-19

Query Match          68.0%; Score 68; DB 10; Length 380;
Best Local Similarity 58.4%; Pred. No. 0.011;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VERVDFTNHLDETRRINK 19
Db 126 VERVDFTNDIOETRRINK 144

RESULT 8
US-60-443-566-4373
; Sequence 4373, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4373
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-4373

Query Match          50.0%; Score 50; DB 14; Length 328;
Best Local Similarity 55.6%; Pred. No. 7.4;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VERVDFTNHLDETRRIN 18
Db 69 VESVDFANAPESRRKIN 86

RESULT 9
US-60-455-444-7455
; Sequence 7455, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
```

```
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7455
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-7455

Query Match          50.0%; Score 50; DB 14; Length 328;
Best Local Similarity 55.6%; Pred. No. 7.4;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VERVDFTNHLDETRRIN 18
Db 69 VESVDFANAPESRRKIN 86

RESULT 10
US-60-465-241-7455
; Sequence 7455, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7455
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-7455

Query Match          50.0%; Score 50; DB 14; Length 328;
Best Local Similarity 55.6%; Pred. No. 7.4;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VERVDFTNHLDETRRIN 18
Db 69 VESVDFANAPESRRKIN 86

RESULT 11
PCT-US02-07215B-12
; Sequence 12, Application PC/TUS0207215B
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard RG
; APPLICANT: Dubose, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-WO
; CURRENT APPLICATION NUMBER: PCT/US02/07215B
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07215B-12

Query Match          50.0%; Score 50; DB 2; Length 390;
Best Local Similarity 55.6%; Pred. No. 9;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VERVDFTNHLDETRRIN 18
Db 131 VESVDFANAPESRRKIN 148
```

```
RESULT 12
PCT-US03-17409-175
; Sequence 175, Application PC/TUS0317409
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: RICKERT, Paula K.
; APPLICANT: KRASNOW, Randi
; TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR LUNG CANCER
; FILE REFERENCE: PA-0051 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/17409
; CURRENT FILING DATE: 2003-06-02
; PRIOR FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PERL Program
; SEQ ID NO 175
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1451328CD1
PCT-US03-17409-175

Query Match          50.0%; Score 50; DB 2; Length 390;
Best Local Similarity 55.6%; Pred. No. 9;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 VERVDFTNHLDETRRNIN 18
      || ||| | |::|: ||
Db      131 VESVDFANAPESRKKIN 148

RESULT 13
US-60-448-285-5
; Sequence 5, Application US/60448285
; GENERAL INFORMATION:
; APPLICANT: ASHTON-RICKARDT, PHILIP
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE INHIBITION OF
; TITLE OF INVENTION: CATHESPINS
; FILE REFERENCE: ARCD:390USP1
; CURRENT APPLICATION NUMBER: US/60/448,285
; CURRENT FILING DATE: 2003-02-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-448-285-5

Query Match          50.0%; Score 50; DB 14; Length 390;
Best Local Similarity 55.6%; Pred. No. 9;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 VERVDFTNHLDETRRNIN 18
      || ||| | |::|: ||
Db      131 VESVDFANAPESRKKIN 148

RESULT 14
US-09-724-676-57391
; Sequence 57391, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentln version 3.2
; SEQ ID NO 57391
```

```
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676-57391

Query Match          50.0%; Score 50; DB 10; Length 410;
Best Local Similarity 55.6%; Pred. No. 9.5;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 VERVDFTNHLDETRRNIN 18
      || ||| | |::|: ||
Db      151 VESVDFANAPESRKKIN 168

RESULT 15
US-09-724-676A-57391
; Sequence 57391, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patentln version 3.2
; SEQ ID NO 57391
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-724-676A-57391

Query Match          50.0%; Score 50; DB 10; Length 410;
Best Local Similarity 55.6%; Pred. No. 9.5;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 VERVDFTNHLDETRRNIN 18
      || ||| | |::|: ||
Db      151 VESVDFANAPESRKKIN 168
```

Search completed: July 11, 2003, 12:35:07
Job time : 19.9545 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 53.5094 Seconds

(Without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-4

Perfect score: 83

Sequence: 1 SYIEVTEGTEATATG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending-Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*

2: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US079_COMB.pep.*

8: /cgn2_6/ptodata/1/paa/US078_COMB.pep.*

9: /cgn2_6/ptodata/1/paa/US077_COMB.pep.*

10: /cgn2_6/ptodata/1/paa/US076_COMB.pep.*

11: /cgn2_6/ptodata/1/paa/US075_COMB.pep.*

12: /cgn2_6/ptodata/1/paa/US074_COMB.pep.*

13: /cgn2_6/ptodata/1/paa/US073_COMB.pep.*

14: /cgn2_6/ptodata/1/paa/US072_COMB.pep.*

15: /cgn2_6/ptodata/1/paa/US071_COMB.pep.*

16: /cgn2_6/ptodata/1/paa/US070_COMB.pep.*

17: /cgn2_6/ptodata/1/paa/US069_COMB.pep.*

18: /cgn2_6/ptodata/1/paa/US068_COMB.pep.*

19: /cgn2_6/ptodata/1/paa/US067_COMB.pep.*

20: /cgn2_6/ptodata/1/paa/US066_COMB.pep.*

21: /cgn2_6/ptodata/1/paa/US065_COMB.pep.*

22: /cgn2_6/ptodata/1/paa/US064_COMB.pep.*

23: /cgn2_6/ptodata/1/paa/US063_COMB.pep.*

24: /cgn2_6/ptodata/1/paa/US062_COMB.pep.*

25: /cgn2_6/ptodata/1/paa/US061_COMB.pep.*

26: /cgn2_6/ptodata/1/paa/US060_COMB.pep.*

27: /cgn2_6/ptodata/1/paa/US059_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	100.0	17	US-09-140-719-4	Sequence 4, Appl1
2	83	100.0	17	US-10-091-442-4	Sequence 4, Appl1
3	83	100.0	118	US-60-188-162-4539	Sequence 4539, Ap
4	83	100.0	380	US-08-091-028A-34	Sequence 34, Appl
5	83	100.0	380	US-09-140-719-34	Sequence 34, Appl
6	83	100.0	380	US-09-508-997A-2	Sequence 2, Appl1

7	83	100.0	380	21	US-09-791-537-281	Sequence 281, App
8	83	100.0	380	23	US-09-936-883A-2	Sequence 2, Appl1
9	83	100.0	380	24	US-10-091-442-34	Sequence 34, Appl
10	83	100.0	380	25	US-10-110-030-2	Sequence 2, Appl1
11	73	88.0	378	21	US-09-791-537-45859	Sequence 45859, A
12	73	88.0	378	24	US-10-023-634-84	Sequence 84, Appl
13	69	83.1	17	4	US-08-091-028A-4	Sequence 4, Appl1
14	69	83.1	88	27	US-60-196-718-4162	Sequence 4162, Ap
15	69	83.1	102	27	US-60-196-718-4163	Sequence 4163, Ap
16	69	83.1	116	22	US-09-834-366-22782	Sequence 22782, A
17	69	83.1	116	27	US-60-197-873-22782	Sequence 22782, A
18	69	83.1	217	1	PCT-US01-09369-134	Sequence 134, App
19	69	83.1	217	1	PCT-US99-01404-8	Sequence 8, Appl1
20	69	83.1	217	16	US-09-735-609-8	Sequence 16, Appl1
21	69	83.1	217	19	US-09-539-330-134	Sequence 134, App
22	69	83.1	217	21	US-09-729-674-134	Sequence 134, App
23	69	83.1	340	21	PCT-US01-01320-169	Sequence 169, App
24	69	83.1	340	21	US-09-764-898-169	Sequence 169, App
25	69	83.1	368	19	US-09-508-997A-6	Sequence 19, Appl1
26	69	83.1	368	23	US-09-936-883A-21	Sequence 21, Appl
27	69	83.1	368	25	US-10-110-030-6	Sequence 6, Appl1
28	69	83.1	379	16	US-09-283-946-2	Sequence 16, App
29	69	83.1	379	22	US-09-852-911-946	Sequence 946, App
30	69	83.1	379	22	US-09-895-814-946	Sequence 946, App
31	69	83.1	379	22	US-10-006-920-946	Sequence 946, App
32	69	83.1	379	24	US-10-006-920-946	Sequence 946, App
33	69	83.1	379	24	US-10-012-896-946	Sequence 946, App
34	69	83.1	379	25	US-10-144-678A-946	Sequence 946, App
35	69	83.1	617	1	PCT-US02-07215-14	Sequence 14, Appl
36	69	83.1	617	22	US-09-852-911-947	Sequence 947, App
37	69	83.1	617	22	US-09-895-814-947	Sequence 947, App
38	69	83.1	617	24	US-10-006-920-947	Sequence 947, App
39	69	83.1	617	24	US-10-012-896-947	Sequence 947, App
40	69	83.1	617	24	US-10-094-944-14	Sequence 14, Appl
41	69	83.1	617	25	US-10-144-678A-947	Sequence 947, Appl
42	68	81.9	379	24	US-10-023-634-85	Sequence 85, Appl
43	68	81.9	379	24	US-10-023-634-85	Sequence 86, Appl
44	68	81.9	388	21	US-09-791-537-127753	Sequence 127753, Sequence 21, Appl
45	67	80.7	135	20	US-09-653-664B-21	

ALIGNMENTS

RESULT 1

US-09-140-719-4

Sequence 4, Application US/09140719

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSUROOKA, Nobuo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, Nobuhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAGUCHI, Kozo

APPLICANT: YAMAGUCHI, Noriomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILED DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-4

Query Match 100.0%; Score 83; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEATATG 17
Db 1 SYIEVTEEGTEATATG 17

RESULT 2

US-10-091-442-4

Sequence 4, Application US/10091442

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IWASA, Fuyuki

TSUROOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhito

KURIHARA, Tatsuya

YAMAICHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-091-442-4

Query Match 100.0%; Score 83; DB 24; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYIEVTEEGTEATATG 17
Db 1 SYIEVTEEGTEATATG 17

RESULT 3

US-60-188-162-4539

Sequence 4539, Application US/60188162

GENERAL INFORMATION:

APPLICANT: Beasley, Ellen

TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

FILE REFERENCE: CLO00518

CURRENT APPLICATION NUMBER: US/60/188,162

CURRENT FILING DATE: 2000-05-09

NUMBER OF SEQ ID NOS: 5094

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4539

LENGTH: 118

TYPE: PRT

ORGANISM: HUMAN

US-60-188-162-4539

OY 1 SYIEVTEEGTEATATG 17
Db 78 SYIEVTEEGTEATATG 94

RESULT 4

US-08-091-028A-34

Sequence 34, Application US/08091028A

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IWASA, Fuyuki

TSUROOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhito

KURIHARA, Tatsuya

APPLICANT:

APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34
Query Match 100.0%; Score 83; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYIEVTEEGTEATATG 17
DB 325 SYIEVTEEGTEATATG 341
RESULT 5
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-2021
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34
Query Match 100.0%; Score 83; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYIEVTEEGTEATATG 17
DB 325 SYIEVTEEGTEATATG 341
RESULT 6
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsln Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2
Query Match 100.0%; Score 83; DB 19; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 SYIEVTEEGTEATATG 17
DB 325 SYIEVTEEGTEATATG 341
RESULT 7

US-09-791-537-281
 ; Sequence 281, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 281
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-791-537-281

Query Match 100.0%; Score 83; DB 21; Length 380;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEVTEGTEATG 17
 |||||||||||||
 Db 325 STIEVTEGTEATG 341

RESULT 8
 US-09-936-883A-2
 ; Sequence 2, Application US/09936883A
 ; GENERAL INFORMATION:
 ; APPLICANT: MIYATA, Toshio
 ; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
 ; FILE REFERENCE: P2-101DIPCT
 ; CURRENT APPLICATION NUMBER: US/09/936,883A
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: JP 1999-75305
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: JP 1999-306623
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-936-883A-2

Query Match 100.0%; Score 83; DB 23; Length 380;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEVTEGTEATG 17
 |||||||||||||
 Db 325 STIEVTEGTEATG 341

RESULT 9
 US-10-091-442-34
 ; Sequence 34, Application US/10091442
 ; GENERAL INFORMATION:
 ; APPLICANT: TSUJIMOTO, Masafumi
 ; INASA, Fuyuki
 ; TSUROOKA, Nobuo
 ; NAKAZATO, Hirosaki
 ; MIURA, Kenju
 ; ISHIDA, Nobuhiko
 ; KURIHARA, Tatsuya
 ; YAMAIICHI, Kozo
 ; YAMAGUCHI, Nozomi
 ; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Burns, Doane, Swecker & Mathis
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/091,442
 FILING DATE: 07-Mar-2002
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/140,719
 FILING DATE: 08-AUG-1998
 APPLICATION NUMBER: US 08/474,661
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/091,028
 FILING DATE: 14-JUL-1993
 APPLICATION NUMBER: JP 4-212305
 FILING DATE: 17-JUL-1992
 APPLICATION NUMBER: JP 6-067339
 FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 380 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 34:
 US-10-091-442-34

Query Match 100.0%; Score 83; DB 24; Length 380;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEVTEGTEATG 17
 |||||||||||||
 Db 325 STIEVTEGTEATG 341

RESULT 10
 US-10-110-030-2
 ; Sequence 2, Application US/10110030
 ; GENERAL INFORMATION:
 ; APPLICANT: MIYATA, Toshio
 ; TITLE OF INVENTION: ANIMAL MODEL FOR MESENCHYMAL PROLIFERATIVE
 ; FILE REFERENCE: SHIM015
 ; CURRENT APPLICATION NUMBER: US/10/110,030
 ; CURRENT FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: 11/285736
 ; PRIOR FILING DATE: 1999-10-06
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-110-030-2

Query Match 100.0%; Score 83; DB 25; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SYEVTGEGTAAATG 17
Db 325 SYEVTGEGTAAATG 341

RESULT 11

US-09-791-537-45859
Sequence 45859, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Blonoxix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 45859
LENGTH: 378
TYPE: PRT
ORGANISM: Bos taurus
US-09-791-537-45859

Query Match 88.0%; Score 73; DB 21; Length 378;
Best Local Similarity 82.4%; Pred. No. 0.0013;

Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYEVTGEGTAAATG 17
Db 321 SYEVTGEGTAAATG 337

RESULT 12

US-10-023-634-84
Sequence 84, Application US/10023634
GENERAL INFORMATION:
APPLICANT: Shinkens, Richard A
APPLICANT: Colman, Steven D
APPLICANT: Spytak, Kimberly A
APPLICANT: Ballinger, Robert A
APPLICANT: Guo, Xiaojia
APPLICANT: Tchernev, Velizar T
APPLICANT: Shenoy, Suresh G
APPLICANT: Li, Li
APPLICANT: Ellerman, Karen
APPLICANT: Zerhusen, Bryan D
APPLICANT: Patlurajan, Meera
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Gusev, Vladimir Y
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomo R
APPLICANT: Gangoli, Esha A
APPLICANT: Malyankar, Uriel M
APPLICANT: Gunther, Erik
APPLICANT: Smltson, Glenda
APPLICANT: Millet, Isabelle
APPLICANT: Gellach, Valerie
TITLE OF INVENTION: Proteins, polynucleotides Encoding them and Methods of
FILE REFERENCE: 21402-221
CURRENT APPLICATION NUMBER: US/10/023.634
CURRENT FILING DATE: 2002-06-28
PRIOR APPLICATION NUMBER: 60/256,025
PRIOR FILING DATE: 2000-12-15
PRIOR APPLICATION NUMBER: 60/265,163
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 60/272,929
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/274,864
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/276,688
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/277,880
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 60/286,409
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/309,246
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/315,600
PRIOR FILING DATE: 2001-08-29
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 84
LENGTH: 378
TYPE: PRT
ORGANISM: Bos taurus
US-10-023-634-84

Query Match 88.0%; Score 73; DB 24; Length 378;
Best Local Similarity 82.4%; Pred. No. 0.0013;

Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SYEVTGEGTAAATG 17
Db 321 SYEVTGEGTAAATG 337

RESULT 13

US-08-091-028-4
Sequence 4, Application US/08091028
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIDRA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEKAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028-4

Query Match 83.1%; Score 69; DB 4; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.0001;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YIEVTEGTEATAATG 17
|||||
Db 2 YIEVTEGTEATAATG 17

RESULT 14
US-08-091-028A-4
Sequence 4, Application US/08091028A

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEX
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028A-4

Query Match 83.1%; Score 69; DB 4; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.0001;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YIEVTEGTEATAATG 17
|||||
Db 2 YIEVTEGTEATAATG 17

RESULT 15
US-60-196-718-4162

Sequence 4162, Application US/60196718

GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: CLO00456
CURRENT APPLICATION NUMBER: US/60/196,718
CURRENT FILING DATE: 2000-04-13
NUMBER OF SEQ ID NOS: 7494
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4162

LENGTH: 88
TYPE: PRT
ORGANISM: HUMAN
US-60-196-718-4162

Query Match 83.1%; Score 69; DB 27; Length 88;
Best Local Similarity 76.5%; Pred. No. 0.00094;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATAATG 17
|||
Db 65 SYIDVSEBSTEATAATG 81

Search completed: July 11, 2003, 12:27:31
Job time : 54.5094 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 15.9751 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-4

Perfect score: 83

Sequence: 1 SYIEVTEGTEATATG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues 1229476

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending-Patents-AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4.*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4.*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4.*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4.*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4.*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4.*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	83	100.0	380	10	US-09-936-883C-2
2	83	100.0	380	12	US-10-126-052A-263
3	83	100.0	380	14	US-60-443-566-3006
4	83	100.0	380	14	US-60-453-444-4692
5	83	100.0	380	14	US-60-465-241-4692
6	69	83.1	368	10	US-09-936-883C-21
7	69	83.1	372	14	US-60-443-566-3717
8	69	83.1	372	14	US-60-455-444-5630
9	69	83.1	372	14	US-60-465-241-5630
10	69	83.1	379	12	US-10-294-025-946
11	69	83.1	392	12	US-10-198-070-62
12	69	83.1	392	12	US-10-198-070-67
13	69	83.1	392	12	US-10-198-070-70
14	69	83.1	617	2	PCT-US02-07215B-14
15	69	83.1	617	12	US-10-294-025-947
16	67	80.7	316	12	US-10-325-745-10
17	67	80.7	339	14	US-60-443-566-4399
18	67	80.7	339	14	US-60-455-444-7456
19	67	80.7	339	14	US-60-465-241-7456

20	67	80.7	391	12	US-10-313-986-110	Sequence 110, App
21	67	80.7	391	14	US-60-443-566-4400	Sequence 4400, Ap
22	67	80.7	391	14	US-60-455-444-7457	Sequence 7457, Ap
23	67	80.7	391	14	US-60-465-241-7457	Sequence 7457, Ap
24	67	80.7	400	12	US-10-313-986-112	Sequence 112, App
25	65	78.3	380	10	US-09-936-883C-19	Sequence 19, App
26	64	77.1	384	2	PCT-US02-07215B-19	Sequence 361, App
27	64	77.1	384	12	US-10-144-779-361	Sequence 199590, .
28	63	75.9	389	12	US-10-424-599-199590	Sequence 199591, .
29	63	75.9	389	12	US-10-424-599-199591	Sequence 199591, .
30	62	74.7	116	12	US-10-130-138A-108	Sequence 108, App
31	62	74.7	188	12	US-10-130-138A-110	Sequence 110, App
32	62	74.7	189	12	US-10-437-963-184620	Sequence 184620, App
33	62	74.7	372	12	US-10-130-138A-150	Sequence 150, App
34	62	74.7	376	14	US-60-448-285-7	Sequence 297, App
35	62	74.7	376	14	US-60-474-733-297	Sequence 358, App
36	62	74.7	377	12	US-10-144-779-358	Sequence 12140, A
37	62	74.7	377	14	US-60-453-135-12140	Sequence 12141, A
38	62	74.7	377	14	US-60-453-135-12141	Sequence 12141, A
39	62	74.7	377	14	US-60-453-050-12140	Sequence 12140, A
40	62	74.7	377	14	US-60-466-412-12141	Sequence 12141, A
41	62	74.7	377	14	US-60-466-412-12141	Sequence 12140, A
42	62	74.7	377	14	US-60-466-412-12141	Sequence 12141, A
43	62	74.7	379	2	PCT-US03-12731-14	Sequence 14, Appl
44	62	74.7	379	2	PCT-US02-07215B-9	Sequence 9, Appl
45	62	74.7	379	14	US-60-448-285-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-936-883C-2
; Sequence 2, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MITSUBISHI, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; TITLE OF INVENTION: Therapeutic
; FILE REFERENCE: P2-101DIPCT
; CURRENT APPLICATION NUMBER: US/09/936, 883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883C-2
Query Match 100.0%; Score 83; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 325 SYIEVTEGTEATATG 17
IIIIIIIIIIIIIIIIIIII
I SYIEVTEGTEATATG 17
RESULTS 2
US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Murray, Natasha
; APPLICANT: Azizi, Richard
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE REFERENCE: 018501-001530US
; CURRENT APPLICATION NUMBER: US/10/126, 052A

```
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/339,245
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 263
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-052A-263
```

```
Query Match
Best Local Similarity 100.0%; Score 83; DB 12; Length 380;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYIEVTEGTEATATG 17
Db 325 SYIEVTEGTEATATG 341
```

```
RESULT 3
US-60-443-566-3006
; Sequence 3006, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3006
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-3006
```

```
Query Match
Best Local Similarity 100.0%; Score 83; DB 14; Length 380;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYIEVTEGTEATATG 17
Db 325 SYIEVTEGTEATATG 341
```

```
RESULT 4
US-60-455-444-4692
; Sequence 4692, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4692
; LENGTH: 380
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-4692
```

```
Query Match
Best Local Similarity 100.0%; Score 83; DB 14; Length 380;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYIEVTEGTEATATG 17
Db 325 SYIEVTEGTEATATG 341
```

```
RESULT 5
US-60-465-241-4692
; Sequence 4692, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 25848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4692
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-4692
```

```
Query Match
Best Local Similarity 100.0%; Score 83; DB 14; Length 380;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYIEVTEGTEATATG 17
Db 325 SYIEVTEGTEATATG 341
```

```
RESULT 6
US-09-936-883C-21
; Sequence 21, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin protein and Use
; FILE REFERENCE: F2-10DIP1CT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-936-883C-21
```

```
Query Match
Best Local Similarity 83.1%; Score 69; DB 10; Length 368;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 SYIEVTEGTEATAT 16
Db 313 SYIEVTEGTEATAT 328
```

```
RESULT 7
```



```
US-60-443-566-3717
; Sequence 3717, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3717
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-3717

Query Match
Best Local Similarity 83.1%; Score 69; DB 14; Length 372;
Best Local Similarity 76.5%; Pred. No. 0.0076;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
   ||:::||||| |||
Db 315 SYLDVSEGTAAATG 331

RESULT 8
US-60-455-444-5630
; Sequence 5630, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5630
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-5630

Query Match
Best Local Similarity 83.1%; Score 69; DB 14; Length 372;
Best Local Similarity 76.5%; Pred. No. 0.0076;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
   ||:::||||| |||
Db 315 SYLDVSEGTAAATG 331

RESULT 9
US-60-465-241-5630
; Sequence 5630, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5630
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-5630
```

```
US-10-294-025-946
; Sequence 946, Application US/10294025
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-946

Query Match
Best Local Similarity 83.1%; Score 69; DB 12; Length 379;
Best Local Similarity 76.5%; Pred. No. 0.0078;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
   ||:::||||| |||
Db 322 SYLDVSEGTAAATG 338

RESULT 10
US-10-294-025-946
; Sequence 946, Application US/10294025
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 946
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-946

Query Match
Best Local Similarity 83.1%; Score 69; DB 12; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0081;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
   ||:::||||| |||
Db 335 SYLDVSEGTAAATG 351

RESULT 11
US-10-198-070-62
; Sequence 62, Application US/10198070
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 62
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-62

Query Match
Best Local Similarity 83.1%; Score 69; DB 12; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0081;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SYIEVTEGTEATATG 17
   ||:::||||| |||
Db 335 SYLDVSEGTAAATG 351
```

RESULT 12
US-10-198-070-67
; Sequence 67, Application US/10198070
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-67

Query Match 83.1%; Score 69; DB 12; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0081;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SYIEVTEGTEATATG 17
||:|:||||| ||||
Db 335 SYLDVSEEGTEAAATG 351

RESULT 13
US-10-198-070-70
; Sequence 70, Application US/10198070.
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 70
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-70

Query Match 83.1%; Score 69; DB 12; Length 392;
Best Local Similarity 76.5%; Pred. No. 0.0081;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SYIEVTEGTEATATG 17
||:|:||||| ||||
Db 335 SYLDVSEEGTEAAATG 351

RESULT 14
PCT-US02-07215B-14

; Sequence 14, Application PC/TUS0207215B
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Clarke, Howard RG
; APPLICANT: Dubose, Robert F
; APPLICANT: Wiley, Steven R
; TITLE OF INVENTION: HUMAN SERPIN POLYPEPTIDES
; FILE REFERENCE: 3223-WO
; CURRENT APPLICATION NUMBER: PCT/US02/07215B
; CURRENT FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (90)..(90)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (215)..(215)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (225)..(225)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (232)..(233)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (289)..(289)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (309)..(309)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (318)..(318)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (322)..(322)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (331)..(332)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (367)..(367)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (369)..(369)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (374)..(374)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (378)..(378)
; OTHER INFORMATION: unsure
; FEATURE: MISC_FEATURE
; NAME/KEY: MISC_FEATURE
; LOCATION: (425)..(425)
; OTHER INFORMATION: unsure
; FEATURE:

NAME/KEY: MISC_FEATURE
LOCATION: (430)..(430)
OTHER INFORMATION: unsure
PCT-US02-07215B-14

Query Match 83.1%; Score 69; DB 2; Length 617;
Best Local Similarity 76.5%; Pred. No. 0.013;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAATG 17
||:::|||||||
Db 560 SYLDVSEGTAAATG 576

RESULT 15
US-10-294-025-947
Sequence 947, Application US/10294025
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolk, John A.
APPLICANT: Kajos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C29
CURRENT APPLICATION NUMBER: US/10/294,025
CURRENT FILING DATE: 2002-11-12
NUMBER OF SEQ ID NOS: 1038
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 947
LENGTH: 617
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: 90, 215, 225, 232, 233, 289, 309, 318, 322, 331, 332, 367,
LOCATION: 369, 374, 378, 425, 430
OTHER INFORMATION: Xaa - Any Amino Acid
US-10-294-025-947

Query Match 83.1%; Score 69; DB 12; Length 617;
Best Local Similarity 76.5%; Pred. No. 0.013;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SYIEVTEGTEATAATG 17
||:::|||||||
Db 560 SYLDVSEGTAAATG 576

Search completed: July 11, 2003, 12:35:07
Job time : 16.0751 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 : Search time 28.3285 Seconds
(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-5

Perfect score: 44
Sequence: 1 QYLRALGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	9	US-09-140-719-5	Sequence 5, Appl1
2	44	100.0	9	US-10-091-442-5	Sequence 5, Appl1
3	44	100.0	78	US-60-160-203-4173	Sequence 4173, Ap
4	44	100.0	118	US-60-188-162-4539	Sequence 4539, Ap
5	44	100.0	380	US-08-091-028A-34	Sequence 34, Appl
6	44	100.0	380	US-09-140-719-34	Sequence 34, Appl

7	44	100.0	380	US-09-508-997A-2	Sequence 2, Appl1
8	44	100.0	380	US-09-791-537-281	Sequence 281, App
9	44	100.0	380	US-09-936-883A-2	Sequence 2, Appl1
10	44	100.0	380	US-10-091-442-34	Sequence 34, Appl
11	44	100.0	380	US-10-110-030-2	Sequence 2, Appl1
12	39	88.6	9	US-08-091-028A-5	Sequence 5, Appl1
13	39	88.6	9	US-08-091-028A-5	Sequence 5, Appl1
14	35	79.5	598	US-09-760-469-1213	Sequence 1213, Ap
15	35	79.5	598	US-10-216-583-1213	Sequence 1213, Ap
16	35	79.5	694	US-09-856-679-12	Sequence 12, Appl
17	35	79.5	694	US-60-118-610-5	Sequence 5, Appl1
18	34	77.3	80	US-09-791-537-48914	Sequence 48914, A
19	34	77.3	210	US-09-791-537-92802	Sequence 92802, A
20	34	77.3	213	PCT-US01-08631-36588	Sequence 36588, A
21	34	77.3	270	US-09-417-507-34770	Sequence 34770, A
22	34	77.3	313	US-09-248-796-17065	Sequence 17065, A
23	34	77.3	313	US-60-096-409-17065	Sequence 17065, A
24	34	77.3	337	US-60-360-039-8700	Sequence 8700, Ap
25	34	77.3	350	US-09-791-537-8066	Sequence 8066, Ap
26	34	77.3	350	US-09-791-537-9963	Sequence 9963, Ap
27	34	77.3	350	US-09-847-208-53	Sequence 53, Appl
28	34	77.3	350	US-09-847-208-53	Sequence 53, Appl
29	34	77.3	409	US-10-179-131-7870	Sequence 7870, Ap
30	34	77.3	434	US-10-179-131-5552	Sequence 5552, Ap
31	34	77.3	597	US-09-540-209B-9878	Sequence 9878, Ap
32	33	75.0	250	US-09-184-082-2	Sequence 2, Appl1
33	33	75.0	537	US-60-360-039-20450	Sequence 20450, A
34	33	75.0	1366	PCT-US98-00615-22	Sequence 22, Appl
35	33	75.0	1366	US-08-781-734-22	Sequence 22, Appl
36	33	75.0	1366	US-08-781-734-22	Sequence 22, Appl
37	33	75.0	1465	US-10-083-357-1310	Sequence 1310, Ap
38	33	75.0	1465	US-60-333-726-1310	Sequence 1310, Ap
39	33	75.0	1890	PCT-US98-00615-88	Sequence 88, Appl
40	32	72.7	80	US-09-791-537-14885	Sequence 14885, Ap
41	32	72.7	92	US-60-195-052-1080	Sequence 1080, Ap
42	32	72.7	92	US-60-195-052-1149	Sequence 1149, Ap
43	32	72.7	96	US-60-195-052-1000	Sequence 1000, Ap
44	32	72.7	97	US-60-213-178-600	Sequence 600, App
45	32	72.7	102	US-60-202-183-189	Sequence 189, App

ALIGNMENTS

RESULT 1
US-09-140-719-5 : Sequence 5, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masaaki
INVENTOR: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
INVENTOR: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
INVENTOR: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
INVENTOR: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-5

Query Match 100.0%; Score 44; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
Db 1 QYLRALGLK 9

RESULT 2
US-10-091-442-5
Sequence 5, Application US/10091442
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IMASA, Fuyuki
TSUBONOKA, Nobuo
NAKAZATO, Hiroshi
MIURA, Kenju
ISHIDA, Nobuhiko
KURIHARA, Tatsuya
YAMAICHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-091-442-5

Query Match 100.0%; Score 44; DB 24; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
Db 1 QYLRALGLK 9

RESULT 3
US-60-160-203-4173
Sequence 4173, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: CL000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4173
LENGTH: 78
TYPE: PRT
ORGANISM: HUMAN
US-60-160-203-4173

Query Match 100.0%; Score 44; DB 27; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRALGLK 9
Db 42 QYLRALGLK 50

RESULT 4
US-60-188-162-4539
Sequence 4539, Application US/60188162
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: CL000518
CURRENT APPLICATION NUMBER: US/60/188,162
CURRENT FILING DATE: 2000-05-09

NUMBER OF SEQ ID NOS: 5094
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4539
LENGTH: 118
TYPE: PRT
ORGANISM: HUMAN
US-60-188-162-4539

Query Match 100.0%; Score 44; DB 27; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGLK 9
db 42 OYLRALGLK 50

RESULT 5
US-08-091-028A-34
Sequence 34, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34
Query Match 100.0%; Score 44; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGLK 9
db 289 OYLRALGLK 297

RESULT 6
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34
Query Match 100.0%; Score 44; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
db 289 OYLRALGLK 297
RESULT 7

US-09-508-997A-2
; Sequence 2, Application US/09508997A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: Megsin Protein
; FILE REFERENCE: SHIM-004
; CURRENT APPLICATION NUMBER: US/09/508,997A
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: JP 09-275302
; PRIOR FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-997A-2
Query Match 100.0%; Score 44; DB 19; Length 380;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 289 OYLRALGLK 297
RESULT 8
US-09-791-537-281
; Sequence 281, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Dancer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 281
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-281
Query Match 100.0%; Score 44; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 289 OYLRALGLK 297
RESULT 9
US-09-936-883A-2
; Sequence 2, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DIPCT
; CURRENT APPLICATION NUMBER: US/09/936,883A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883A-2
Query Match 100.0%; Score 44; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYLRALGLK 9
DB 289 OYLRALGLK 297
RESULT 10
US-10-091-442-34
; Sequence 34, Application US/10091442
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, Nobuhito
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozi
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm R.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34
Query Match 100.0%; Score 44; DB 24; Length 380;

Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGLK 9
Db 289 OYLRALGLK 297

RESULT 11

US-10-110-030-2
Sequence 2, Application US/10110030
GENERAL INFORMATION:
APPLICANT: Miyata, Toshio
TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
FILE REFERENCE: SHIM015
CURRENT APPLICATION NUMBER: US/10/110.030
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 11/285736
PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-10-110-030-2

Query Match 100.0%; Score 44; DB 25; Length 380;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGLK 9
Db 289 OYLRALGLK 297

RESULT 12

US-08-091-028-5
Sequence 5, Application US/08091028
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091.028
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feuyl, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028-5

Query Match 88.6%; Score 39; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4; 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
Db 2 YLRALGLK 9

RESULT 13

US-08-091-028A-5
Sequence 5, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091.028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028A-5

Query Match
Best Local Similarity 88.6%; Score 39; DB 4; Length 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YLRALGL 9
Db 2 YLRALGL 9

RESULT 14
US-09-760-469-1213
Sequence 1213, Application US/09760469
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT254
CURRENT APPLICATION NUMBER: US/09/760,469
CURRENT FILING DATE: 2001-01-16
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1983
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1213
LENGTH: 598
TYPE: PRT
ORGANISM: Homo sapiens
US-09-760-469-1213

Query Match
Best Local Similarity 79.5%; Score 35; DB 21; Length 598;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGL 8
Db 455 RYLRALGL 462

RESULT 15
US-10-216-583-1213
Sequence 1213, Application US/10216583
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT254C1N
CURRENT APPLICATION NUMBER: US/10/216,583
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 09/760,469
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1983

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1213
LENGTH: 598
TYPE: PRT
ORGANISM: Homo sapiens
US-10-216-583-1213

Query Match
Best Local Similarity 79.5%; Score 35; DB 26; Length 598;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLRALGL 8
Db 455 RYLRALGL 462

Search completed: July 11, 2003, 12:27:33
Job time: 30.3285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 8.45738 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-5
Perfect score: 44
Sequence: 1 OYLKALGLK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues

Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID.	Description
1	44	100.0	380	10 US-09-936-883C-2	Sequence 2, Appl 1
2	44	100.0	380	12 US-10-126-052A-263	Sequence 263, App
3	44	100.0	380	14 US-60-443-566-3006	Sequence 3006, App
4	44	100.0	380	14 US-60-455-444-4692	Sequence 4692, App
5	44	100.0	380	14 US-60-465-241-4692	Sequence 4692, App
6	36	81.8	730	12 US-10-437-963-150051	Sequence 150051, App
7	35	79.5	1293	12 US-10-437-963-117663	Sequence 117663, App
8	34	77.3	337	12 US-10-369-493-8700	Sequence 8700, App
9	33	75.0	282	12 US-10-282-122A-47577	Sequence 47577, App
10	33	75.0	537	12 US-10-369-493-20450	Sequence 20450, App
11	33	75.0	582	12 US-10-437-963-115178	Sequence 115178, App
12	32	72.7	124	12 US-10-425-114-60805	Sequence 60805, App
13	32	72.7	127	12 US-10-425-114-59319	Sequence 59319, App
14	32	72.7	150	12 US-10-425-114-60833	Sequence 60833, App
15	32	72.7	159	12 US-10-425-114-71125	Sequence 71125, App
16	32	72.7	172	11 US-10-213-573-10778	Sequence 10778, App
17	32	72.7	176	12 US-10-219-051B-10334	Sequence 10334, App
18	32	72.7	180	12 US-10-425-114-62321	Sequence 62321, App
19	32	72.7	190	12 US-10-425-114-66361	Sequence 66361, App

	20	32	72.7	228	2	PCT-US03-01605-318	Sequence 318, App
	21	32	72.7	257	12	US-10-437-963-115387	Sequence 115387, App
	22	32	72.7	281	12	US-10-030-019A-380	Sequence 380, App
	23	32	72.7	342	12	US-10-282-122A-60837	Sequence 60837, App
	24	32	72.7	377	12	US-10-437-963-198078	Sequence 198078, App
	25	32	72.7	382	12	US-10-282-122A-48213	Sequence 48213, App
	26	32	72.7	408	12	US-10-369-493-6303	Sequence 6303, App
	27	32	72.7	447	10	US-09-103-611D-248	Sequence 248, App
	28	32	72.7	456	10	US-09-800-770-15	Sequence 15, Appl
	29	32	72.7	465	12	US-10-366-683-18919	Sequence 18919, App
	30	32	72.7	465	12	US-10-419-128-18919	Sequence 18919, App
	31	32	72.7	467	12	US-10-380-731-743	Sequence 743, App
	32	32	72.7	473	12	US-10-446-203-8988	Sequence 8988, App
	33	32	72.7	481	12	US-10-380-731-742	Sequence 742, App
	34	32	72.7	481	14	US-60-443-566-4405	Sequence 4405, App
	35	32	72.7	481	14	US-60-452-680-22239	Sequence 22239, App
	36	32	72.7	481	14	US-60-453-135-13699	Sequence 13699, App
	37	32	72.7	481	14	US-60-453-050-13699	Sequence 13699, App
	38	32	72.7	481	14	US-60-455-444-7275	Sequence 7275, App
	39	32	72.7	481	14	US-60-465-241-7275	Sequence 7275, App
	40	32	72.7	481	14	US-60-466-412-13659	Sequence 13659, App
	41	32	72.7	523	12	US-10-282-122A-77204	Sequence 77204, App
	42	32	72.7	591	2	PCT-US03-07858-702	Sequence 702, App
	43	32	72.7	591	2	PCT-US03-07858-2393	Sequence 2393, App
	44	32	72.7	591	2	PCT-US03-07858A-702	Sequence 702, App
	45	32	72.7	591	2	PCT-US03-07858A-2393	Sequence 2393, App

ALIGNMENTS

RESULT 1
US-09-936-883C-2
; Sequence 2, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DIPCT
; CURRENT APPLICATION NUMBER: US/09/936, 883C
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PPT
; ORGANISM: Homo sapiens
; US-09-936-883C-2

Query Match 100.0%; Score 44; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYLKALGLK 9
DB 289 OYLKALGLK 297

RESULT 2
US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Richard
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer. Compositions and
; FILE REFERENCE: 018501-00153005
; CURRENT APPLICATION NUMBER: US/10/126, 052A

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-455-444-4692
Query Match
Best Local Similarity 100.0%; Score 44; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYRALGLK 9
Db 289 OYRALGLK 297
RESULT 5
US-60-465-241-4692
; Sequence 4692, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-052A-263

Query Match
Best Local Similarity 100.0%; Score 44; DB 12; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYRALGLK 9
Db 289 OYRALGLK 297
RESULT 3
US-60-443-566-3006
; Sequence 3006, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3006
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-3006

Query Match
Best Local Similarity 100.0%; Score 44; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OYRALGLK 9
Db 289 OYRALGLK 297
RESULT 4
US-60-455-444-4692
; Sequence 4692, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-455-444-4692

Query Match
Best Local Similarity 100.0%; Score 44; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYRALGLK 9
Db 289 OYRALGLK 297

RESULT 5
US-60-465-241-4692
; Sequence 4692, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-4692

Query Match
Best Local Similarity 100.0%; Score 44; DB 14; Length 380;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OYRALGLK 9
Db 289 OYRALGLK 297

RESULT 6
US-10-437-963-150051
; Sequence 150051, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 150051
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50323C.1.pep
US-10-437-963-150051

Query Match
Best Local Similarity 81.8%; Score 36; DB 12; Length 730;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 OYRALGLK 9
Db 289 OYRALGLK 297

Db 524 QYLRAIGOK 532

RESULT 7

US-10-437-963-117663

; Sequence 117663, Application US/10437963

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(5322)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 117663

; LENGTH: 1293

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_21048C.1.pcp

US-10-437-963-117663

Query Match

Best Local Similarity 79.5%; Score 35; DB 12; Length 1293;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYLRAIGLK 9

Db 593 RYLKALGIR 601

RESULT 8

US-10-369-493-8700

; Sequence 8700, Application US/10369493

; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 8700

; LENGTH: 337

; TYPE: PRT

; ORGANISM: Ralstonia metallidurans

US-10-369-493-8700

Query Match

Best Local Similarity 77.3%; Score 34; DB 12; Length 337;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYLRAIGLK 9

Db 182 QYARAGLR 190

RESULT 9

US-10-282-122A-47577

; Sequence 47577, Application US/10282122A

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: EITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See file wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 47577

; LENGTH: 282

; TYPE: PRT

; ORGANISM: Burkholderia cepacia

US-10-282-122A-47577

Query Match

Best Local Similarity 75.0%; Score 33; DB 12; Length 282;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYLRAIGLK 9

Db 38 QYLKSTGLR 46

RESULT 10

US-10-369-493-20450

; Sequence 20450, Application US/10369493

; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 20450

LENGTH: 537
TYPE: PRT
ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20450

Query Match 75.0%; Score 33; DB 12; Length 537;
Best Local Similarity 87.5%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 YLRALGLK 9
DB 81 YLRALGLK 88

RESULT 11
US-10-437-963-115178
; Sequence 115178, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Mei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 115178
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_187C.1.pep
US-10-437-963-115178

Query Match 75.0%; Score 33; DB 12; Length 582;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 QYLRALGLK 9
DB 356 QYLRALGLK 364

RESULT 12
US-10-425-114-60805
; Sequence 60805, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60805
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-048-F4_FLI.pep
US-10-425-114-60805

Query Match 72.7%; Score 32; DB 12; Length 124;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRALGLK 9
DB 111 LRALGLK 117

RESULT 13
US-10-425-114-59319
; Sequence 59319, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59319
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB189-027-D3_FLI.pep
US-10-425-114-59319

Query Match 72.7%; Score 32; DB 12; Length 127;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRALGLK 9
DB 114 LRALGLK 120

RESULT 14
US-10-425-114-60833
; Sequence 60833, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60833
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-105-E8_FLI.pep
US-10-425-114-60833

Query Match 72.7%; Score 32; DB 12; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRALGLK 9
DB 137 LRALGLK 143

RESULT 15

US-10-425-114-71125
 ; Sequence 71125, Application US/10425114
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 71125
 ; LENGTH: 159
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3597-011-F8_FLI.pep
 US-10-425-114-71125

Query Match 72.7%; Score 32; DB 12; Length 159;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRALGLK 9
 |||||
 Db 114 LRALGLK 120

Search completed: July 11, 2003, 12:35:09
 Job time : 10.5574 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 62.9522 Seconds

(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-6

Perfect score: 98

Sequence: 1 ADLSGRASGRLYSRMGK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US006.COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US006.COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US007.COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US008.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	20	15	US-09-140-719-6
2	98	100.0	20	24	US-10-091-442-6
3	92	93.9	20	4	US-08-091-028-6
4	92	93.9	20	4	US-08-091-028-6
5	90	91.8	78	27	US-60-160-203-4173
6	90	91.8	118	27	US-60-188-162-4539

7	90	91.8	380	4	US-08-091-028A-34	Sequence 34, Appl
8	90	91.8	380	15	US-09-140-719-34	Sequence 34, Appl
9	90	91.8	380	19	US-09-508-997A-2	Sequence 2, Appl
10	90	91.8	380	21	US-09-791-537-281	Sequence 281, App
11	90	91.8	380	23	US-09-936-883A-2	Sequence 2, Appl
12	90	91.8	380	24	US-10-091-442-34	Sequence 34, Appl
13	90	91.8	380	25	US-10-110-030-2	Sequence 2, Appl
14	83	84.7	368	19	US-09-508-997A-6	Sequence 6, Appl
15	83	84.7	368	23	US-09-936-883A-21	Sequence 21, Appl
16	83	84.7	368	25	US-10-110-030-6	Sequence 6, Appl
17	83	84.7	380	19	US-09-508-997A-4	Sequence 4, Appl
18	83	84.7	380	23	US-09-936-883A-19	Sequence 19, Appl
19	83	84.7	380	25	US-10-110-030-4	Sequence 4, Appl
20	54	55.1	195	21	US-09-791-537-102306	Sequence 102306,
21	53	54.1	405	21	US-09-791-537-99016	Sequence 99016, A
22	51	52.0	125	20	US-09-617-682A-3038	Sequence 3038, Ap
23	51	52.0	210	21	US-09-708-427-53024	Sequence 53024, A
24	51	52.0	210	26	US-10-219-999-47784	Sequence 47784, A
25	51	52.0	210	26	US-10-219-999-62585	Sequence 62585, A
26	51	52.0	210	27	US-60-312-544-5910	Sequence 9910, Ap
27	51	52.0	210	27	US-60-324-109-28769	Sequence 28769, A
28	51	52.0	306	21	US-09-708-427-53023	Sequence 53023, A
29	51	52.0	306	26	US-10-219-999-54875	Sequence 54875, A
30	51	52.0	306	26	US-09-791-537-45859	Sequence 45859, A
31	51	52.0	378	24	US-10-023-634-84	Sequence 84, Appl
32	51	52.0	378	24	US-10-029-386-12104	Sequence 32104, A
33	50	51.0	615	1	PCT-US01-04942A-356	Sequence 356, App
34	50	51.0	615	20	US-09-629-469A-12768	Sequence 12768, A
35	50	51.0	615	21	US-09-791-537-112135	Sequence 112135,
36	50	51.0	615	26	US-10-221-278-356	Sequence 356, App
37	48	49.0	423	24	US-10-023-634-84	Sequence 83, Appl
38	48	49.0	423	21	US-09-791-537-132657	Sequence 132657,
39	48	49.0	591	21	US-60-186-718-6010	Sequence 6010, Ap
40	47	48.0	94	27	US-60-186-718-6011	Sequence 6011, Ap
41	47	48.0	94	27	US-10-023-634-20	Sequence 20, Appl
42	47	48.0	365	24	PCT-US01-13360A-59	Sequence 59, Appl
43	47	48.0	394	1	PCT-US01-13360A-59	Sequence 59, Appl
44	47	48.0	394	1	PCT-US01-13360A-59	Sequence 59, Appl
45	47	48.0	395	22	US-09-823-187-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-140-719-6
; Sequence 6, Application US/09140719
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURUOKA, Nobuo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, Nobuhiko
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22113-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/140.719

;; FILING DATE: 08-AUG-1998
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/474,661
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-247
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-09-140-719-6

Query Match 100.0%; Score 98; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMCK 20
Db 1 ADLSGIASGRLYISRMCK 20

RESULT 2
US-10-091-442-6
; Sequence 6, Application US/10091442
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; IMASA, Fuyuki
; TSURUOKA, Nobuo
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, Nobuhiro
; KURIHARA, Tatsuya
; YAMAGUCHI, Kozo
; YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998

;; APPLICATION NUMBER: US 08/474,661
;; FILING DATE: 07-JUN-1995
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; APPLICATION NUMBER: JP 6-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McGowan, Malcolm K.
;; REGISTRATION NUMBER: 39,300
;; REFERENCE/DOCKET NUMBER: 001560-247
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-2021
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
;; US-10-091-442-6

Query Match 100.0%; Score 98; DB 24; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMCK 20
Db 1 ADLSGIASGRLYISRMCK 20

RESULT 3
US-08-091-028-6
; Sequence 6, Application US/08091028
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; IMASA, Fuyuki
; TSURUOKA, Nobuo
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, Nobuhiro
; KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/091,028
; FILING DATE: 14-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E

REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-091-028-6

Query Match 93.9%; Score 92; DB 4; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXMK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 ADLSGIASGRLYISRMXMK 20

RESULT 4
US-08-091-028A-6
Sequence 6, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASHI, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-091-028A-6

Query Match 93.9%; Score 92; DB 4; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXMK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 ADLSGIASGRLYISRMXMK 20

RESULT 5
US-60-160-203-4173
Sequence 4173, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: CLO00116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4173
LENGTH: 78
TYPE: PRT
ORGANISM: HUMAN
US-60-160-203-4173

Query Match 91.8%; Score 90; DB 27; Length 78;
Best Local Similarity 95.0%; Pred. No. 1.5e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXMK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 58 ADLSGIASGRLYISRMXMK 77

RESULT 6
US-60-188-162-4539
Sequence 4539, Application US/60188162
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: CLO00518
CURRENT APPLICATION NUMBER: US/60/188,162
CURRENT FILING DATE: 2000-05-09
NUMBER OF SEQ ID NOS: 5094
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4539
LENGTH: 118
TYPE: PRT
ORGANISM: HUMAN
US-60-188-162-4539

Query Match 91.8%; Score 90; DB 27; Length 118;
Best Local Similarity 95.0%; Pred. No. 2.4e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMXMK 20
| | | | | | | | | | | | | | | | | | | | | |
Db 58 ADLSGIASGRLYISRMXMK 77

RESULT 7
US-08-091-028A-34
Sequence 34, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANKE
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34

Query Match 91.8%; Score 90; DB 4; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSGIASGRLYISRMKG 20
DB 305 ADLSGIASGRLYISRMKG 324

RESULT 8
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 91.8%; Score 90; DB 15; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSGIASGRLYISRMKG 20
DB 305 ADLSGIASGRLYISRMKG 324

RESULT 9
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsln Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2

Query Match 91.8%; Score 90; DB 19; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY      1 ADLSCIAGGRLYISRMKGK 20
DB      305 ADLSCIAGGRLYISRMHMK 324

RESULT 10
US-09-791-537-281
; Sequence 281, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonometry, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 281
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-281

Query Match          91.8%; Score 90; DB 21; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 ADLSCIAGGRLYISRMKGK 20
DB      305 ADLSCIAGGRLYISRMHMK 324

RESULT 11
US-09-936-883A-2
; Sequence 2, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DPIPT
; CURRENT APPLICATION NUMBER: US/09/936,883A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883A-2

Query Match          91.8%; Score 90; DB 23; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 ADLSCIAGGRLYISRMKGK 20
DB      305 ADLSCIAGGRLYISRMHMK 324

RESULT 12
US-10-091-442-34
; Sequence 34, Application US/10091442
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; TSUROUOKA, Nobuo
; NAKAZATO, Hiroshi

MIURA, Kenju
ISHIDA, Nobuhito
KURIHARA, Tatsuya
YAMACHI, Kozo
YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match          91.8%; Score 90; DB 24; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 ADLSCIAGGRLYISRMKGK 20
DB      305 ADLSCIAGGRLYISRMHMK 324

RESULT 13
US-10-110-030-2
; Sequence 2, Application US/10110030
; GENERAL INFORMATION:
; APPLICANT: Miyata, Toshio
; TITLE OF INVENTION: ANIMAL MODEL FOR MESENCHYMAL PROLIFERATIVE
; FILE REFERENCE: SHIM015
; CURRENT APPLICATION NUMBER: US/10/110,030
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 11/285736
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-10-110-030-2

Query Match 91.8%; Score 90; DB 25; Length 380;
Best Local Similarity 95.0%; Pred. No. 9.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMGMK 20
|||||
Db 305 ADLSGIASGRLYISRMGMK 324

RESULT 14
US-09-508-997A-6
Sequence 6, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 368
TYPE: PRT
ORGANISM: Mus musculus
US-09-508-997A-6

Query Match 84.7%; Score 83; DB 19; Length 368;
Best Local Similarity 80.0%; Pred. No. 0.00013;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMGMK 20
|||||
Db 293 ADLSGIASGRLYISRMGMK 312

RESULT 15
US-09-936-883A-21
Sequence 21, Application US/09936883A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
OF INVENTION: thereof
FILE REFERENCE: P2-101DPICT
CURRENT APPLICATION NUMBER: US/09/936,883A
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 368
TYPE: PRT
ORGANISM: Mus musculus
US-09-936-883A-21

Query Match 84.7%; Score 83; DB 23; Length 368;
Best Local Similarity 80.0%; Pred. No. 0.00013;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMGMK 20
|||||
Db 293 ADLSGIASGRLYISRMGMK 312

Search completed: July 11, 2003, 12:27:34
Job time : 63.9522 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 18.7942 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-6
Perfect score: 98
Sequence: 1 ADLSGIASGRLYISRMNGK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues
Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	91.8	380	10	US-09-936-883C-2
2	90	91.8	380	12	US-10-126-052A-263
3	90	91.8	380	14	US-60-443-566-3006
4	90	91.8	380	14	US-60-455-444-4692
5	90	91.8	380	14	US-60-465-241-4692
6	83	84.7	368	10	US-09-936-883C-21
7	83	84.7	380	10	US-09-936-883C-19
8	51	52.0	233	12	US-10-425-114-57272
9	51	52.0	233	12	US-10-425-114-68267
10	51	52.0	306	12	US-10-425-114-59731
11	51	52.0	306	12	US-10-425-114-59936
12	47	48.0	394	12	US-10-258-951-59
13	47	48.0	405	2	PCT-US02-38526-416
14	47	48.0	405	14	US-60-443-566-4407
15	47	48.0	405	14	US-60-455-444-7462
16	47	48.0	405	14	US-60-465-241-7462
17	47	48.0	425	2	PCT-US02-38526-408
18	47	48.0	425	2	PCT-US02-07215B-2
19	47	48.0	425	12	US-10-419-277-2

20	47	48.0	425	14	US-60-443-566-4406	Sequence 4406, Ap
21	47	48.0	425	14	US-60-455-444-7461	Sequence 7461, Ap
22	47	48.0	425	14	US-60-465-241-7461	Sequence 7461, Ap
23	47	48.0	617	2	PCT-US02-38526-415	Sequence 415, App
24	46	46.9	257	2	PCT-US02-40225-3054	Sequence 3054, Ap
25	46	46.9	257	12	US-10-320-797-3054	Sequence 3054, Ap
26	46	46.9	377	12	US-10-144-779-358	Sequence 358, App
27	45	45.9	238	10	US-09-724-676-67567	Sequence 67567, A
28	45	45.9	238	10	US-09-724-676-67568	Sequence 67568, A
29	45	45.9	238	10	US-09-724-676-67569	Sequence 67569, A
30	45	45.9	238	10	US-09-724-676-67570	Sequence 67570, A
31	45	45.9	238	10	US-09-724-676-67571	Sequence 67571, A
32	45	45.9	238	10	US-09-724-676-67572	Sequence 67572, A
33	45	45.9	238	10	US-09-724-676-67573	Sequence 67573, A
34	45	45.9	238	10	US-09-724-676-67574	Sequence 67574, A
35	45	45.9	238	10	US-09-724-676-67575	Sequence 67575, A
36	45	45.9	238	10	US-09-724-676-67576	Sequence 67576, A
37	45	45.9	238	10	US-09-724-676-67577	Sequence 67577, A
38	45	45.9	238	10	US-09-724-676-67578	Sequence 67578, A
39	45	45.9	238	10	US-09-724-676-67579	Sequence 67579, A
40	45	45.9	238	10	US-09-724-676-67580	Sequence 67580, A
41	45	45.9	238	10	US-09-724-676-67581	Sequence 67581, A
42	45	45.9	238	10	US-09-724-676-67582	Sequence 67582, A
43	45	45.9	238	10	US-09-724-676-67583	Sequence 67583, A
44	45	45.9	238	10	US-09-724-676-67584	Sequence 67584, A
45	45	45.9	238	10	US-09-724-676-67585	Sequence 67585, A

ALIGNMENTS

RESULT 1
US-09-936-883C-2
; Sequence 2, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: P2-10DIPRCT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-936-883C-2

Query Match 91.8%; Score 90; DB 10; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADLSGIASGRLYISRMNGK 20
DB 305 ADLSGIASGRLYISRMNGK 324

RESULT 2
US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE REFERENCE: 018501-001530US
; CURRENT APPLICATION NUMBER: US/10/126,052A

```

: CURRENT FILING DATE: 2002-04-18
: PRIOR APPLICATION NUMBER: US 60/284,770
: PRIOR FILING DATE: 2001-04-18
: PRIOR APPLICATION NUMBER: US 60/290,492
: PRIOR FILING DATE: 2001-05-10
: PRIOR APPLICATION NUMBER: US 60/339,245
: PRIOR FILING DATE: 2001-11-09
: PRIOR APPLICATION NUMBER: US 60/350,666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/334,370
: PRIOR FILING DATE: 2001-11-29
: PRIOR APPLICATION NUMBER: US 60/372,246
: PRIOR FILING DATE: 2002-04-12
: NUMBER OF SEQ ID NOS: 691
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 263
: LENGTH: 380
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-126-052A-263
```

```

Query Match          91.8%; Score 90; DB 12; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 ADLSGIASGRLYISRMGK 20
Db      305 ADLSGIASGRLYISRMHK 324
```

```

RESULT 3
US-60-443-566-3006
: Sequence 3006, Application US/60443566
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: FILE REFERENCE: C1001447
: CURRENT APPLICATION NUMBER: US/60/443,566
: CURRENT FILING DATE: 2003-01-30
: NUMBER OF SEQ ID NOS: 25102
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3006
: LENGTH: 380
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-443-566-3006
```

```

Query Match          91.8%; Score 90; DB 14; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 ADLSGIASGRLYISRMGK 20
Db      305 ADLSGIASGRLYISRMHK 324
```

```

RESULT 4
US-60-445-444-4692
: Sequence 4692, Application US/60455444
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: FILE REFERENCE: C1001455
: CURRENT APPLICATION NUMBER: US/60/445,444
: CURRENT FILING DATE: 2003-03-18
: NUMBER OF SEQ ID NOS: 50986
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4692
: LENGTH: 380
```

```

: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-455-444-4692
```

```

Query Match          91.8%; Score 90; DB 14; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 ADLSGIASGRLYISRMGK 20
Db      305 ADLSGIASGRLYISRMHK 324
```

```

RESULT 5
US-60-465-241-4692
: Sequence 4692, Application US/60465241
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: FILE REFERENCE: C1001468
: CURRENT APPLICATION NUMBER: US/60/465,241
: CURRENT FILING DATE: 2003-04-23
: NUMBER OF SEQ ID NOS: 258418
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4692
: LENGTH: 380
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-465-241-4692
```

```

Query Match          91.8%; Score 90; DB 14; Length 380;
Best Local Similarity 95.0%; Pred. No. 1.3e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 ADLSGIASGRLYISRMGK 20
Db      305 ADLSGIASGRLYISRMHK 324
```

```

RESULT 6
US-09-936-883C-21
: Sequence 21, Application US/09936883C
: GENERAL INFORMATION:
: APPLICANT: MIYATA, Toshio
: TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
: FILE REFERENCE: P2-101DIPCT
: CURRENT APPLICATION NUMBER: US/09/936,883C
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: JP 1999-75305
: PRIOR FILING DATE: 1999-03-19
: PRIOR APPLICATION NUMBER: JP 1999-306623
: PRIOR FILING DATE: 1999-10-28
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 21
: LENGTH: 368
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-936-883C-21
```

```

Query Match          84.7%; Score 83; DB 10; Length 368;
Best Local Similarity 80.0%; Pred. No. 1.9e-05;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      1 ADLSGIASGRLYISRMGK 20
Db      293 ADLSGIASGRLYVSKIMHK 312
```

```

RESULT 7
```



```
US-09-936-883C-19
: Sequence 19, Application US/09936883C
: GENERAL INFORMATION:
: APPLICANT: MIYATA, Toshio
: TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
: FILE REFERENCE: P2-101DPCT
: CURRENT APPLICATION NUMBER: US/09/936, 883C
: PRIOR APPLICATION NUMBER: JP 1999-75305
: PRIOR FILING DATE: 1999-03-19
: PRIOR APPLICATION NUMBER: JP 1999-306623
: PRIOR FILING DATE: 1999-10-28
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 19
: LENGTH: 380
: TYPE: PRT
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 51, 94
: OTHER INFORMATION: Xaa is unknown.
US-09-936-883C-19

Query Match      84.7%: Score 83; DB 10; Length 380;
Best Local Similarity 80.0%: Pred. No. 2e-05;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      1 ADLSGIASGRLYISRMGR 20
Db      305 ADLSGIASGRLYVSKMKR 324

RESULT 8
US-10-425-114-57272
: Sequence 57272, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovacic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(5313)B
: CURRENT APPLICATION NUMBER: US/10/425, 114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 57272
: LENGTH: 233
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-ZMFLB73109H11_FLI.pep
US-10-425-114-57272

Query Match      52.0%: Score 51; DB 12; Length 233;
Best Local Similarity 64.7%: Pred. No. 2.9;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      1 ADLSGIASGRLYISRM 17
Db      27 AGLSGIESGRLYISNL 43

RESULT 9
US-10-425-114-68267
: Sequence 68267, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovacic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(5313)B
: CURRENT APPLICATION NUMBER: US/10/425, 114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 68267
: LENGTH: 233
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-ZMFLMO17295D05_FLI.pep
US-10-425-114-68267

Query Match      52.0%: Score 51; DB 12; Length 233;
Best Local Similarity 64.7%: Pred. No. 2.9;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      1 ADLSGIASGRLYISRM 17
Db      27 AGLSGIESGRLYISNL 43

RESULT 10
US-10-425-114-59731
: Sequence 59731, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovacic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(5313)B
: CURRENT APPLICATION NUMBER: US/10/425, 114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 59731
: LENGTH: 306
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: LIB3957-001-A2_FLI.pep
US-10-425-114-59731

Query Match      52.0%: Score 51; DB 12; Length 306;
Best Local Similarity 64.7%: Pred. No. 4;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY      1 ADLSGIASGRLYISRM 17
Db      100 AGLSGIESGRLYISNL 116

RESULT 11
US-10-425-114-59936
: Sequence 59936, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovacic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(5313)B
```

;; CURRENT APPLICATION NUMBER: US/10/425,114
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 73128
;; SEQ ID NO 59936
;; LENGTH: 306
;; TYPE: PRT
;; ORGANISM: Zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: LIB3689-229-C3_Flt.pep
US-10-425-114-59936

Query Match 52.0%; Score 51; DB 12; Length 306;
Best Local Similarity 64.7%; Pred. No. 4;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMK 17
Db 100 AGLSGIESGTYKLTISNL 116

RESULT 12
US-10-258-951-59
;; Sequence 59, Application US/10258951
;; GENERAL INFORMATION:
;; APPLICANT: Agarwal, Pankaj
;; APPLICANT: Murock, Paul R.
;; APPLICANT: Rizvi, Safia K.
;; APPLICANT: Smith, Randall F.
;; APPLICANT: Xiang, Zhaoying
;; APPLICANT: Kadinick, Karen
;; APPLICANT: Lai, Ying-Ta
;; APPLICANT: Xie, Qing
;; TITLE OF INVENTION: NOVEL COMPOUNDS
;; FILE REFERENCE: GP50025
;; CURRENT APPLICATION NUMBER: US/10/258,951
;; PRIOR FILING DATE: 2002-10-28
;; PRIOR APPLICATION NUMBER: PCT/US01/13360
;; PRIOR FILING DATE: 2001-04-26
;; PRIOR APPLICATION NUMBER: 60/199,963
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/203,336
;; PRIOR FILING DATE: 2000-05-11
;; PRIOR APPLICATION NUMBER: 60/207,087
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/207,546
;; PRIOR FILING DATE: 2000-05-26
;; NUMBER OF SEQ ID NOS: 78
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 59
;; LENGTH: 394
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-258-951-59

Query Match 48.0%; Score 47; DB 12; Length 394;
Best Local Similarity 45.0%; Pred. No. 26;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMK 20
Db 317 ADLTGISPSNLTLSKIHK 336

RESULT 13
PCT-US02-38526-416
;; Sequence 416, Application PC/TUS0238526
;; GENERAL INFORMATION:
;; APPLICANT: Ghosh, Malabika
;; APPLICANT: Tang, Y. Tom
;; APPLICANT: Wang, Jian-Rui
;; APPLICANT: Wang, Zhiwei
;; APPLICANT: Zhao, Qing
;; APPLICANT: Xu, Chongjun

;; APPLICANT: Mulero, Julio J
;; APPLICANT: Boyle, Bryan J.
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL POLYPEPTIDES AND POLYN
;; FILE REFERENCE: HYS-B1/PCT
;; CURRENT APPLICATION NUMBER: PCT/US02/38526
;; CURRENT FILING DATE: 2002-12-03
;; PRIOR APPLICATION NUMBER: US 09/488,725
;; PRIOR FILING DATE: 2000-01-21
;; PRIOR APPLICATION NUMBER: US 09/552,317
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: PCT/US00/35017
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: US 09/620,312
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: PCT/US00/34263
;; PRIOR FILING DATE: 2000-12-26
;; PRIOR APPLICATION NUMBER: US 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; PRIOR APPLICATION NUMBER: US 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: PCT/US01/03800
;; PRIOR FILING DATE: 2001-02-05
;; PRIOR APPLICATION NUMBER: US 09/598,075
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: PCT/US01/04098
;; PRIOR FILING DATE: 2001-02-05
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 653
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 416
;; LENGTH: 405
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US02-38526-416

Query Match 48.0%; Score 47; DB 2; Length 405;
Best Local Similarity 45.0%; Pred. No. 27;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMK 20
Db 328 ADLTGISPSNLTLSKIHK 347

RESULT 14
US-60-443-566-4407
;; Sequence 4407, Application US/60443566
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: BEGOVICH, Ann
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; FILE REFERENCE: CL001447
;; CURRENT APPLICATION NUMBER: US/60/443,566
;; CURRENT FILING DATE: 2003-01-30
;; NUMBER OF SEQ ID NOS: 25102
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4407
;; LENGTH: 405
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-443-566-4407

Query Match 48.0%; Score 47; DB 14; Length 405;
Best Local Similarity 45.0%; Pred. No. 27;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGIASGRLYISRMK 20
Db 328 ADLTGISPSNLTLSKIHK 347

RESULT 15

US-60-455-444-7462
: Sequence 7462, Application US/60455444
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01455
: CURRENT APPLICATION NUMBER: US/60/455,444
: CURRENT FILING DATE: 2003-03-18
: NUMBER OF SEQ ID NOS: 50986
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7462
: LENGTH: 405
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-455-444-7462

Query Match 48.0%; Score 47; DB 14; Length 405;
Best Local Similarity 45.0%; Pred. No. 27;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 ADLSGIASGGRLYISRMGK 20
|||:|: ||:::|
Db 328 ADLTGISPPNLYLXKTIHK 347

Search completed: July 11, 2003, 12:35:10
Job time : 19.8942 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 15.738 Seconds

(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-7

Perfect score: 26

Sequence: 1 LYDAK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US111_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US112_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	26	100.0	5	US-08-091-028-7
2	26	100.0	4	US-08-091-028A-7
3	26	100.0	5	US-09-140-719-7
4	26	100.0	15	US-10-091-442-7
5	26	100.0	41	US-60-196-718-3839
6	26	100.0	58	US-09-727-892-99

7	26	100.0	58	US-09-727-892A-99	Sequence 99, Appl
8	26	100.0	75	US-60-160-209-2854	Sequence 2854, Ap
9	26	100.0	75	US-60-163-062-871	Sequence 871, App
10	26	100.0	76	US-09-758-471-3660	Sequence 3660, Ap
11	26	100.0	76	US-10-235-953-3660	Sequence 3660, Ap
12	26	100.0	81	US-09-540-209B-7383	Sequence 7383, Ap
13	26	100.0	99	US-60-160-203-3968	Sequence 3968, Ap
14	26	100.0	99	US-60-163-123-1299	Sequence 1299, Ap
15	26	100.0	114	US-60-162-247-3259	Sequence 3259, Ap
16	26	100.0	117	US-09-791-537-71302	Sequence 71302, A
17	26	100.0	118	PCT-US00-07661-169	Sequence 169, App
18	26	100.0	118	PCT-US00-07661-171	Sequence 171, App
19	26	100.0	118	PCT-US00-07661-169	Sequence 169, App
20	26	100.0	118	PCT-US00-07661-171	Sequence 171, App
21	26	100.0	133	US-10-015-127-9753	Sequence 9753, Ap
22	26	100.0	131	US-09-513-996A-33705	Sequence 33705, A
23	26	100.0	151	US-09-688-051-892	Sequence 892, App
24	26	100.0	155	US-09-513-996A-33704	Sequence 33704, A
25	26	100.0	155	US-09-688-051-891	Sequence 891, App
26	26	100.0	167	US-09-614-450-5014	Sequence 5014, Ap
27	26	100.0	171	US-09-614-450-5013	Sequence 5013, Ap
28	26	100.0	196	US-08-827-356-2912	Sequence 2912, Ap
29	26	100.0	199	US-09-791-537-147392	Sequence 147392, A
30	26	100.0	209	US-60-360-039-17822	Sequence 17822, A
31	26	100.0	211	US-09-675-784A-7278	Sequence 7278, Ap
32	26	100.0	212	PCT-US01-08631-55189	Sequence 55189, A
33	26	100.0	214	US-60-140-956-1585	Sequence 1585, Ap
34	26	100.0	215	US-60-167-217-8763	Sequence 8763, Ap
35	26	100.0	216	US-09-540-209B-5974	Sequence 5974, Ap
36	26	100.0	221	US-09-770-767-57481	Sequence 57481, A
37	26	100.0	221	US-09-370-849B-182166	Sequence 182166, A
38	26	100.0	227	US-09-611-529-4520	Sequence 4520, Ap
39	26	100.0	263	US-09-513-996A-29703	Sequence 29703, A
40	26	100.0	263	US-09-513-996A-46566	Sequence 46566, A
41	26	100.0	263	US-09-689-980-341	Sequence 341, App
42	26	100.0	281	US-09-791-537-115423	Sequence 115423, A
43	26	100.0	282	US-09-739-449-8924	Sequence 8924, Ap
44	26	100.0	282	US-09-803-110-8924	Sequence 8924, Ap
45	26	100.0	291	US-10-219-999-47475	Sequence 47475, A

ALIGNMENTS

RESULT 1
US-08-091-028-7
Sequence 7, Application US/08091028
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
INVENTOR: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Koze
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028

FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028-7

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 1 LYDAK 5

RESULT 2

US-08-091-028A-7
Sequence 7, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028A-7

Query Match 100.0%; Score 26; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 1 LYDAK 5

RESULT 3

US-09-140-719-7
Sequence 7, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-7

Query Match 100.0%; Score 26; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
|||||
DB 1 LYDAK 5

RESULT 4
US-10-091-442-7

; Sequence 7, Application US/10091442
; GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

TSUROOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhito

KURIHARA, Tatsuya

YAMAGUCHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-Aug-1998

APPLICATION NUMBER: US 08/474,661

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-247

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-091-442-7

Query Match 100.0%; Score 26; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
|||||
DB 1 LYDAK 5

RESULT 5
US-60-196-718-3839

; Sequence 3839, Application US/60196718
; GENERAL INFORMATION:

APPLICANT: Bonazzi, Vivien

TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: C1000456

CURRENT APPLICATION NUMBER: US/60/196,718

CURRENT FILING DATE: 2000-04-13

NUMBER OF SEQ ID NOS: 7494

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3839

LENGTH: 41

TYPE: PRT

ORGANISM: HUMAN

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(41)

OTHER INFORMATION: Xaa = Any Amino Acid

US-60-196-718-3839

Query Match 100.0%; Score 26; DB 27; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
|||||
DB 10 LYDAK 14

RESULT 6
US-09-727-892-99

; Sequence 99, Application US/09727892
; GENERAL INFORMATION:

APPLICANT: PhageTech, Inc.

APPLICANT: PELLETIER, Jerry

APPLICANT: GROS, Philippe

APPLICANT: DUBOW, Michael

TITLE OF INVENTION: DNA SEQUENCES FROM STAPHYLOCOCCUS AUREUS BACTERIOPHAGE 44 AHJD

TITLE OF INVENTION: THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES

FILE REFERENCE: 073406-0302

CURRENT APPLICATION NUMBER: US/09/727,892

CURRENT FILING DATE: 2000-12-01

NUMBER OF SEQ ID NOS: 159

SOFTWARE: Patentin version 3.0

SEQ ID NO 99

LENGTH: 58

TYPE: PRT

ORGANISM: Staphylococcus aureus Bacteriophage 44 AHJD

US-09-727-892-99

Query Match 100.0%; Score 26; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
|||||
DB 29 LYDAK 33

```
RESULT 7
US-09-727-892A-99
; Sequence 99, Application US/09727892A
; GENERAL INFORMATION:
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GROS, PHILIPPE
; APPLICANT: DUBOW, MICHAEL
; TITLE OF INVENTION: DNA SEQUENCES FROM STAPHYLOCOCCUS AUREUS BACTERIOPHAGE 44AHJD
; FILE REFERENCE: 073406/0302
; CURRENT APPLICATION NUMBER: US/09/727, 892A
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/168, 777
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 99
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Staphylococcus aureus Bacteriophage 44AHJD
US-09-727-892A-99

Query Match
Best Local Similarity 100.0%; Score 26; DB 21; Length 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 29 LYDAK 33

RESULT 8
US-60-160-209-2854
; Sequence 2854, Application US/60160209
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: C1000113
; CURRENT APPLICATION NUMBER: US/60/160, 209
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4646
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2854
; LENGTH: 75
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-209-2854

Query Match
Best Local Similarity 100.0%; Score 26; DB 27; Length 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 33 LYDAK 37

RESULT 9
US-60-163-062-871
; Sequence 871, Application US/60163062
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: C1000134
; CURRENT APPLICATION NUMBER: US/60/163, 062
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1302
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 871
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Human
US-60-163-062-871

Query Match
Best Local Similarity 100.0%; Score 26; DB 27; Length 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 33 LYDAK 37

RESULT 10
US-09-758-471-3660
; Sequence 3660, Application US/09758471
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH002
; CURRENT APPLICATION NUMBER: US/09/758, 471
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179, 065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 5386
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3660
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-471-3660

Query Match
Best Local Similarity 100.0%; Score 26; DB 21; Length 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 72 LYDAK 76

RESULT 11
US-10-235-953-3660
; Sequence 3660, Application US/10235953
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH002CIN
; CURRENT APPLICATION NUMBER: US/10/235, 953
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/758, 471
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179, 065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 5386
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3660
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-235-953-3660

Query Match
Best Local Similarity 100.0%; Score 26; DB 26; Length 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
```


Db 72 LYDAK 76

RESULT 12
US-09-540-209B-7383

; Sequence 7383, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540.209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7383
; LENGTH: 81
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-7383

Query Match 100.0%; Score 26; DB 19; Length 81;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 69 LYDAK 73

RESULT 13
US-60-160-203-3968

; Sequence 3968, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVLEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160.203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3968
; LENGTH: 99
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-3968

Query Match 100.0%; Score 26; DB 27; Length 99;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 50 LYDAK 54

RESULT 14
US-60-163-123-1299

; Sequence 1299, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVLEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000137
; CURRENT APPLICATION NUMBER: US/60/163.123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1299
; LENGTH: 99

; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1299

Query Match 100.0%; Score 26; DB 27; Length 99;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 50 LYDAK 54

RESULT 15
US-60-162-247-3259

; Sequence 3259, Application US/60162247
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVLEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000127
; CURRENT APPLICATION NUMBER: US/60/162.247
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 5442
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3259
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Human
US-60-162-247-3259

Query Match 100.0%; Score 26; DB 27; Length 114;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 65 LYDAK 69

Search completed: July 11, 2003, 12:27:36
Job time : 17.738 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 4.69854 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-7
Perfect score: 26
Sequence: 1 LYDAK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues
Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA.New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	26	100.0	209	12	US-10-369-493-17822
2	26	100.0	211	12	US-10-09-675-784A-7278
3	26	100.0	213	12	US-10-437-963-115384
4	26	100.0	227	10	US-09-950-084-4520
5	26	100.0	253	12	US-10-282-122A-60526
6	26	100.0	257	12	US-10-424-599-219617
7	26	100.0	265	12	US-10-437-963-144912
8	26	100.0	291	12	US-10-425-114-55804
9	26	100.0	319	12	US-10-434-437-98
10	26	100.0	333	10	US-09-724-676-56008
11	26	100.0	343	12	US-09-724-676A-56008
12	26	100.0	345	12	US-10-366-683-21042
13	26	100.0	345	12	US-10-419-128-21042
14	26	100.0	380	12	US-09-936-883C-2
15	26	100.0	380	12	US-10-126-052A-263
16	26	100.0	380	14	US-60-443-566-3006
17	26	100.0	380	14	US-60-455-444-4692
18	26	100.0	380	14	US-60-465-241-4692
19	26	100.0	395	12	US-10-282-122A-48743

20	26	100.0	415	10	US-09-134-000C-3595	Sequence 3595, Ap
21	26	100.0	415	10	US-09-134-000C-3595	Sequence 3595, Ap
22	26	100.0	415	12	US-10-434-663-3595	Sequence 3595, Ap
23	26	100.0	427	12	US-10-454-437-96	Sequence 96, Appl
24	26	100.0	430	12	US-10-156-761-14800	Sequence 14800, A
25	26	100.0	454	12	US-10-310-154-660	Sequence 660, App
26	26	100.0	458	12	US-10-446-203-11632	Sequence 11632, A
27	26	100.0	489	2	PCT-US02-38445-42	Sequence 42, Appl
28	26	100.0	491	12	US-10-417-886-9542	Sequence 9542, Ap
29	26	100.0	496	12	US-10-072-012-771	Sequence 771, App
30	26	100.0	516	12	US-10-437-963-190810	Sequence 190810, .
31	26	100.0	521	12	US-10-072-012-769	Sequence 769, App
32	26	100.0	524	12	US-10-369-493-20966	Sequence 20966, A
33	26	100.0	529	12	US-10-431-652-6406	Sequence 6406, Ap
34	26	100.0	558	12	US-10-369-493-10997	Sequence 10997, A
35	26	100.0	559	12	US-10-431-652-5522	Sequence 5522, Ap
36	26	100.0	567	12	US-10-156-761-7781	Sequence 7781, Ap
37	26	100.0	574	12	US-10-369-493-23086	Sequence 23086, A
38	26	100.0	587	12	US-10-369-493-19378	Sequence 19378, A
39	26	100.0	588	12	US-10-366-683-24185	Sequence 24185, A
40	26	100.0	588	12	US-10-419-128-24185	Sequence 24185, A
41	26	100.0	597	12	US-10-282-122A-62084	Sequence 62084, A
42	26	100.0	625	12	US-10-425-114-58457	Sequence 58457, A
43	26	100.0	692	12	US-10-282-122A-72270	Sequence 72270, A
44	26	100.0	707	12	US-10-156-761-7899	Sequence 7899, Ap
45	26	100.0	712	12	US-10-156-761-15026	Sequence 15026, A

ALIGNMENTS

RESULT 1
US-10-369-493-17822
; Sequence 17822, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17822
; LENGTH: 209
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-17822

Query Match 100.0%; Score 26; DB 12; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LYDAK 5
Db 128 LYDAK 132

RESULT 2
US-09-675-784A-7278
; Sequence 7278, Application US/09675784A
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORK
; APPLICANT: ZENG, QIANPONG

```

; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: ASPERGILLUS FUMIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
; FILE REFERENCE: 2976-4020051
; CURRENT APPLICATION NUMBER: US/09/675,784A
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/156,338
; NUMBER OF SEQ ID NOS: 13925
; SEQ ID NO 7278
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-675-784A-7278

Query Match
Best Local Similarity 100.0%; Score 26; DB 10; Length 211;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
   |||||
Db 206 LYDAK 210

RESULT 3
US-10-437-963-115384
; Sequence 115384, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbezuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 115384
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18987C.1.pcp
US-10-437-963-115384

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 213;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
   |||||
Db 67 LYDAK 71

RESULT 4
US-09-950-084-4520
; Sequence 4520, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
```

```

; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 4520
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-4520

Query Match
Best Local Similarity 100.0%; Score 26; DB 10; Length 227;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
   |||||
Db 128 LYDAK 132

RESULT 5
US-10-282-122A-60526
; Sequence 60526, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes In Microorganisms
; FILE REFERENCE: ELITRA.03A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
```

PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60526
LENGTH: 253
TYPE: PRT
ORGANISM: Listeria monocytogenes
US-10-282-122A-60526

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 253;
Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 231 LYDAK 235

RESULT 6
US-10-424-599-219617
Sequence 219617, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285864
SEQ ID NO 219617
LENGTH: 257
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_40342C.1.pep
US-10-424-599-219617

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 257;
Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 239 LYDAK 243

RESULT 7
US-10-437-963-144912
Sequence 144912, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbaszuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 144912
LENGTH: 265
TYPE: PRT
ORGANISM: Oryza sativa

FEATURE:
NAME/KEY: unsure
LOCATION: (1).(265)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_45682C.1.pep
US-10-437-963-144912

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 265;
Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 45 LYDAK 49

RESULT 8
US-10-425-114-55804
Sequence 55804, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 55804
LENGTH: 291
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GMROPIC033E03_FLI.pep
US-10-425-114-55804

Query Match
Best Local Similarity 100.0%; Score 26; DB 12; Length 291;
Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYDAK 5
Db 34 LYDAK 38

RESULT 9
US-10-454-437-98
Sequence 98, Application US/10454437
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
FILE REFERENCE: BG1-128CPN
CURRENT APPLICATION NUMBER: US/10/454,437
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932127.2

;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932128.0
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932129.9
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: DE 19932226.0
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19932920.6
;; PRIOR FILING DATE: 1999-07-14
;; PRIOR APPLICATION NUMBER: DE 19932922.2
;; PRIOR FILING DATE: 1999-07-14
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 442
;; SEQ ID NO 98
;; LENGTH: 319
;; TYPE: PRT
;; ORGANISM: Corynebacterium glutamicum
US-10-454-437-98

Query Match 100.0%; Score 26; DB 12; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 246 LYDAK 250

RESULT 10
US-09-724-676-56008
;; Sequence 56008, Application US/09724676
;; GENERAL INFORMATION:
;; APPLICANT: Compugen LTD
;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 56008
;; LENGTH: 333
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-724-676-56008

Query Match 100.0%; Score 26; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 33 LYDAK 37

RESULT 11
US-09-724-676A-56008
;; Sequence 56008, Application US/09724676A
;; GENERAL INFORMATION:
;; APPLICANT: Compugen LTD
;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676A
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 56008
;; LENGTH: 333
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-724-676A-56008

Query Match 100.0%; Score 26; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LYDAK 5
DB 33 LYDAK 37

RESULT 12
US-10-366-683-21042
;; Sequence 21042, Application US/10366683
;; GENERAL INFORMATION:
;; APPLICANT: Rubenfield, Marc J.
;; APPLICANT: Nolling, Jork
;; APPLICANT: Deloughery, Craig
;; APPLICANT: Bush, David
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; FILE REFERENCE: PATH03-04
;; CURRENT APPLICATION NUMBER: US/10/366,683
;; CURRENT FILING DATE: 2003-02-13
;; PRIOR APPLICATION NUMBER: 09/252,991
;; PRIOR FILING DATE: 1999-02-18
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 21042
;; LENGTH: 345
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-21042

Query Match 100.0%; Score 26; DB 12; Length 345;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 331 LYDAK 335

RESULT 13
US-10-419-128-21042
;; Sequence 21042, Application US/10419128
;; GENERAL INFORMATION:
;; APPLICANT: Marc J. Rubenfield et al.
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/10/419,128
;; CURRENT FILING DATE: 2003-04-21
;; PRIOR APPLICATION NUMBER: US/09/252,991
;; PRIOR FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 21042
;; LENGTH: 345
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-21042

Query Match 100.0%; Score 26; DB 12; Length 345;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYDAK 5
DB 331 LYDAK 335

RESULT 14
US-09-936-883C-2
;; Sequence 2, Application US/09936883C

```
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: P2-101DPICT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-936-883C-2
```

```
Query Match          100.0%; Score 26; DB 10; Length 380;
Best Local Similarity 100.0%; Pred No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 LYDAK 5
        |||||
Db       121 LYDAK 125
```

```
RESULT 15
US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Richard
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE REFERENCE: 018501-001530US
; CURRENT APPLICATION NUMBER: US/10/126,052A
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/339,245
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 263
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-126-052A-263
```

```
Query Match          100.0%; Score 26; DB 12; Length 380;
Best Local Similarity 100.0%; Pred No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 LYDAK 5
        |||||
Db       121 LYDAK 125
```

Search completed: July 11, 2003, 12:35:11
Job time : 5.79854 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 15.738 Seconds

(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-8

Perfect score: 28

Sequence: 1 NTEWK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main.*
1: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US089.COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US091.COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US097.COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US099.COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US102.COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US103.COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US104.COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US105.COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US106.COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US107.COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US108.COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US109.COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US110.COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US111.COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US112.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	5	4	US-08-091-028-8
2	28	100.0	5	4	US-08-091-028A-8
3	28	100.0	5	15	US-09-140-719-8
4	28	100.0	5	24	US-09-091-442-8
5	28	100.0	78	27	US-60-160-203-4173
6	28	100.0	118	27	US-60-188-163-4539

7	28	100.0	126	20	US-09-621-976-5429	Sequence 5429, Ap
8	28	100.0	145	17	US-60-147-499-5429	Sequence 5429, Ap
9	28	100.0	126	27	US-09-328-352-6374	Sequence 6374, Ap
10	28	100.0	163	21	US-09-731-872-445	Sequence 445, Ap
11	28	100.0	163	22	US-09-876-997-445	Sequence 445, Ap
12	28	100.0	163	27	US-60-169-629-607	Sequence 607, Ap
13	28	100.0	163	27	US-60-187-470-607	Sequence 607, Ap
14	28	100.0	181	20	US-09-675-784A-12333	Sequence 12333, A
15	28	100.0	214	16	US-09-252-691-7915	Sequence 7915, Ap
16	28	100.0	214	16	US-09-252-691C-7915	Sequence 7915, Ap
17	28	100.0	278	21	US-09-760-446A-1177	Sequence 1177, Ap
18	28	100.0	278	26	US-10-206-664-1177	Sequence 1177, Ap
19	28	100.0	294	25	US-10-108-260A-4567	Sequence 4567, Ap
20	28	100.0	298	18	US-09-417-507-40692	Sequence 40692, A
21	28	100.0	380	4	US-08-091-028A-34	Sequence 34, Appl
22	28	100.0	380	15	US-09-140-719-34	Sequence 34, Appl
23	28	100.0	380	19	US-09-508-997A-2	Sequence 2, Appl
24	28	100.0	380	21	US-09-791-537-281	Sequence 281, App
25	28	100.0	380	23	US-09-936-883A-2	Sequence 2, Appl
26	28	100.0	380	24	US-10-091-442-34	Sequence 34, Appl
27	28	100.0	380	25	US-10-110-030-2	Sequence 2, Appl
28	28	100.0	399	24	US-10-029-180-64	Sequence 64, Appl
29	28	100.0	400	16	US-09-248-796-19018	Sequence 19018, A
30	28	100.0	400	27	US-60-096-409-19018	Sequence 19018, A
31	28	100.0	417	21	US-09-760-446A-1893	Sequence 1893, Ap
32	28	100.0	417	26	US-10-206-664-1893	Sequence 1893, Ap
33	28	100.0	420	19	US-09-509-234A-41	Sequence 41, Appl
34	28	100.0	420	19	US-09-509-234C-41	Sequence 41, Appl
35	28	100.0	420	21	US-09-791-537-79163	Sequence 79163, A
36	28	100.0	420	21	US-09-791-537-132920	Sequence 132920, A
37	28	100.0	428	19	US-09-513-996A-37385	Sequence 37385, A
38	28	100.0	428	21	US-09-708-427-10543	Sequence 10543, A
39	28	100.0	441	19	US-09-513-996A-37384	Sequence 37384, A
40	28	100.0	441	19	US-09-708-427-10542	Sequence 10542, A
41	28	100.0	479	21	US-09-513-996A-37383	Sequence 37383, A
42	28	100.0	479	21	US-09-708-427-10541	Sequence 10541, A
43	28	100.0	562	21	US-09-791-537-22455	Sequence 22455, A
44	28	100.0	584	20	US-09-604-957-6	Sequence 6, Appl
45	28	100.0	584	23	US-09-995-746A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-091-028-8
Sequence 8, Application US/08091028
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028

FILED DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028-8

Query Match 100.0%; Score 28; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 1 NYEMK 5

RESULT 2
US-08-091-028A-8
Sequence 8, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22131-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANEK
REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028A-8

Query Match 100.0%; Score 28; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 1 NYEMK 5

RESULT 3
US-09-140-719-8
Sequence 8, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22131-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-8

Query Match 100.0%; Score 28; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 1 NYEMK 5

RESULT 4

US-10-091-442-8

Sequence 8, Application US/10091442

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

TSURUDOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhiko

KURIHARA, Tatsuya

YAMAUCHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-247

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-091-442-8

Query Match 100.0%; Score 28; DB 24; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 1 NYEMK 5

RESULT 5

US-60-160-203-4173

Sequence 4173, Application US/60160203

GENERAL INFORMATION:

APPLICANT: BONAZZI, VIVIAN

TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: C1000116

CURRENT APPLICATION NUMBER: US/60/160,203

CURRENT FILING DATE: 1999-10-19

NUMBER OF SEQ ID NOS: 6374

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4173

LENGTH: 78

TYPE: PRT

ORGANISM: HUMAN

US-60-160-203-4173

Query Match 100.0%; Score 28; DB 27; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 37 NYEMK 41

RESULT 6

US-60-188-162-4539

Sequence 4539, Application US/60188162

GENERAL INFORMATION:

APPLICANT: Beasley, Ellen

TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: C1000518

CURRENT APPLICATION NUMBER: US/60/188,162

CURRENT FILING DATE: 2000-05-09

NUMBER OF SEQ ID NOS: 5094

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4539

LENGTH: 118

TYPE: PRT

ORGANISM: HUMAN

US-60-188-162-4539

Query Match 100.0%; Score 28; DB 27; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
Db 37 NYEMK 41

RESULT 7

US-09-621-976-5429

Sequence 5429, Application US/09621976

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J. B.

APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5429
LENGTH: 126
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -41...-1
NAME/KEY: UNSURE
LOCATION: 62
OTHER INFORMATION: Xaa - Lys,Arg
US-09-621-976-5429

Query Match
Best Local Similarity 100.0%; Score 28; DB 20; Length 126;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
|||||
DB 122 NTEMK 126

RESULT 8
US-60-147-499-5429
Sequence 5429, Application US/60147499
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/60/147,499
CURRENT FILING DATE: 1999-08-05
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5429
LENGTH: 126
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -41...-1
FEATURE:
NAME/KEY: UNSURE
LOCATION: 62
OTHER INFORMATION: Xaa - Lys,Arg
US-60-147-499-5429

Query Match
Best Local Similarity 100.0%; Score 28; DB 27; Length 126;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
|||||
DB 122 NTEMK 126

RESULT 9
US-09-328-352-6374
Sequence 6374, Application US/09328352
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: CTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6374
LENGTH: 145
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6374

Query Match
Best Local Similarity 100.0%; Score 28; DB 17; Length 145;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
|||||
DB 28 NTEMK 32

RESULT 10
US-09-731-872-445
Sequence 445, Application US/09731872
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US3.REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 445
LENGTH: 163
TYPE: PRT
ORGANISM: Homo sapiens
US-09-731-872-445

Query Match
Best Local Similarity 100.0%; Score 28; DB 21; Length 163;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NTEMK 5
|||||
DB 159 NTEMK 163

RESULT 11
US-09-876-997-445
Sequence 445, Application US/09876997
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US4.CIP
CURRENT APPLICATION NUMBER: US/09/876,997
CURRENT FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: US 09/731,872
PRIOR FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 445
LENGTH: 163
TYPE: PRT
ORGANISM: Homo sapiens
US-09-876-997-445

Query Match 100.0%; Score 28; DB 22; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 159 NYEMK 163

RESULT 12

US-60-169-629-607
; Sequence 607, Application US/60169629
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: CDNAS for Secreted Proteins
; FILE REFERENCE: GENSET.071PRF
; CURRENT APPLICATION NUMBER: US/60/169,629
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 607
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -77..-1
US-60-169-629-607

Query Match 100.0%; Score 28; DB 27; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 159 NYEMK 163

RESULT 13

US-60-187-470-607
; Sequence 607, Application US/60187470
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: CDNAS for Secreted Proteins
; FILE REFERENCE: 78.US2.PRO
; CURRENT APPLICATION NUMBER: US/60/187,470
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 715
; SOFTWARE: Patent.pm
; SEQ ID NO 607
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -77..-1
US-60-187-470-607

Query Match 100.0%; Score 28; DB 27; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 159 NYEMK 163

RESULT 14

US-09-675-784A-12333
; Sequence 12333, Application US/09675784A
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORR
; APPLICANT: ZENG, QIANDONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: ASPERGILLUS FUMIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
; AND USES THEREFOR
; FILE REFERENCE: 2976-4020US1
; CURRENT APPLICATION NUMBER: US/09/675,784A
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/156,338
; PRIOR FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 13925
; SEQ ID NO 12333
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-675-784A-12333

Query Match 100.0%; Score 28; DB 20; Length 181;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 92 NYEMK 96

RESULT 15
US-09-252-691-7915
; Sequence 7915, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTE
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691B
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 7915
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-09-252-691-7915

Query Match 100.0%; Score 28; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
Db 18 NYEMK 22

Search completed: July 11, 2003, 12:27:38
Job time: 17.738 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 4.69854 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-8

Perfect score: 28
Sequence: 1 NYEMK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues

Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US05_NEW_COMB.pep4:*
4: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	28	100.0	116	10	US-09-724-676-95964
2	28	100.0	116	10	US-09-724-676-95964
3	28	100.0	135	12	US-10-424-599-265931
4	28	100.0	145	12	US-10-431-652-6374
5	28	100.0	158	10	US-09-724-676-95968
6	28	100.0	158	10	US-09-724-676-95968
7	28	100.0	175	12	US-10-424-599-254001
8	28	100.0	181	10	US-09-675-784A-12333
9	28	100.0	205	12	US-10-424-599-265929
10	28	100.0	214	12	US-10-417-886-7915
11	28	100.0	380	10	US-09-936-883C-2
12	28	100.0	380	12	US-10-126-052A-263
13	28	100.0	380	14	US-60-443-566-3006
14	28	100.0	380	14	US-60-455-444-4692
15	28	100.0	380	14	US-60-465-241-4692
16	28	100.0	819	14	US-60-443-566-2783
17	28	100.0	819	14	US-60-452-680-17072
18	28	100.0	819	14	US-60-453-135-10384
19	28	100.0	819	14	US-60-453-050-10384

20	28	100.0	819	14	US-60-466-412-10384	Sequence 10384, A
21	28	100.0	847	2	PCT-US02-19669A-498	Sequence 498, App
22	25	89.3	32	12	US-10-019-596A-3	Sequence 3, App11
23	25	89.3	32	12	US-10-019-596A-5	Sequence 5, App11
24	25	89.3	32	12	US-10-019-596A-6	Sequence 6, App11
25	25	89.3	32	12	US-10-019-596A-7	Sequence 7, App11
26	25	89.3	32	12	US-10-019-596A-8	Sequence 8, App11
27	25	89.3	32	12	US-10-019-596A-11	Sequence 11, App1
28	25	89.3	37	10	US-10-019-596A-35	Sequence 35, App1
29	25	89.3	43	10	US-09-950-083B-3405	Sequence 3405, App
30	25	89.3	52	12	US-10-424-599-234765	Sequence 234765, App
31	25	89.3	55	12	US-10-424-599-191002	Sequence 191002, App
32	25	89.3	59	12	US-10-424-599-232746	Sequence 232746, App
33	25	89.3	61	10	US-09-724-676-93757	Sequence 93757, A
34	25	89.3	61	10	US-09-724-676-93757	Sequence 93757, A
35	25	89.3	61	12	US-10-424-599-162592	Sequence 162592, A
36	25	89.3	67	12	US-10-424-599-211936	Sequence 211936, App
37	25	89.3	72	10	US-09-724-676-93759	Sequence 93759, A
38	25	89.3	72	10	US-09-724-676-93759	Sequence 93759, A
39	25	89.3	74	10	US-09-513-999C-7069	Sequence 7069, App
40	25	89.3	74	10	US-09-513-999C-7444	Sequence 7444, App
41	25	89.3	90	10	US-09-513-999C-7444	Sequence 7444, App
42	25	89.3	90	10	US-09-513-999C-7436	Sequence 7436, App
43	25	89.3	101	10	US-09-864-408A-7436	Sequence 7436, App
44	25	89.3	101	12	US-10-437-963-186021	Sequence 186021, App
45	25	89.3	136	12	US-10-437-963-184199	Sequence 184199, App

ALIGNMENTS

RESULT 1
US-09-724-676-95964
; Sequence 95964, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 95964
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-95964
Query Match 100.0%; Score 28; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
LENGTH: 116
NYEMK 5
Db 6 NYEMK 10
RESULT 2
US-09-724-676A-95964
; Sequence 95964, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 95964
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676A-95964

Query Match 100.0%; Score 28; DB 10; Length 116;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
DB 6 NYEMK 10

RESULT 3
US-10-424-599-265931

; Sequence 265931, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265931
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82156C.1.pep
US-10-424-599-265931

Query Match 100.0%; Score 28; DB 12; Length 135;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
DB 112 NYEMK 116

RESULT 4
US-10-431-652-6374
; Sequence 6374, Application US/10431652
; GENERAL INFORMATION:
; APPLICANT: Breton, Gary L.
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PAT03-08
; CURRENT APPLICATION NUMBER: US/10/431,652
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: US 09/328,352
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 60/088,701
; PRIOR FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6374
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-431-652-6374

Query Match 100.0%; Score 28; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
DB 28 NYEMK 32

RESULT 5

US-09-724-676-95968
; Sequence 95968, Application US/09724676

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 95968
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-95968

Query Match 100.0%; Score 28; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
DB 6 NYEMK 10

RESULT 6
US-09-724-676A-95968

; Sequence 95968, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 95968
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-95968

Query Match 100.0%; Score 28; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYEMK 5
|||||
DB 6 NYEMK 10

RESULT 7
US-10-424-599-254001

; Sequence 254001, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254001
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71386C.1.pep
US-10-424-599-254001

Query Match 100.0%; Score 28; DB 12; Length 175;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Query Match 100.0%; Score 28; DB 12; Length 175;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 14 NYEMK 18

RESULT 8

US-09-675-784A-12333
; Sequence 12333, Application US/09675784A
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORR
; APPLICANT: ZENG, QIANDONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: ASPERGILLUS FUMIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: 2976-4020051
; CURRENT APPLICATION NUMBER: US/09/675,784A
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/156,338
; PRIOR FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 13925
; SEQ ID NO 12333
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-675-784A-12333

Query Match 100.0%; Score 28; DB 10; Length 181;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 92 NYEMK 96

RESULT 9

US-10-424-599-265929
; Sequence 265929, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265929
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82154C.1.pep
US-10-424-599-265929

Query Match 100.0%; Score 28; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 98 NYEMK 102

RESULT 10

US-10-417-886-7915
; Sequence 7915, Application US/10417886
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/10/417,886
; CURRENT FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: US/09/252,691C
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/094,145
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 60/074,787
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 11326
; SEQ ID NO 7915
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-417-886-7915

Query Match 100.0%; Score 28; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 18 NYEMK 22

RESULT 11

US-09-936-883C-2
; Sequence 2, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: F2-101DIPCT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883C-2

Query Match 100.0%; Score 28; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEMK 5
DB 284 NYEMK 288

RESULT 12

US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Murray, Richard
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of lung cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer

```

; FILE REFERENCE: 018501-001530US
; CURRENT APPLICATION NUMBER: US/10/126,052A
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/339,245
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 263
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-052A-263

```

```

Query Match          100.0%; Score 28; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 NYEMK 5
        |||||
Db      284 NYEMK 288

```

```

RESULT 13
US-60-443-566-3006
; Sequence 3006, Application US/60/443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3006
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-3006

```

```

Query Match          100.0%; Score 28; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 NYEMK 5
        |||||
Db      284 NYEMK 288

```

```

RESULT 14
US-60-455-444-4692
; Sequence 4692, Application US/60/455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 4692
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-4692

```

```

Query Match          100.0%; Score 28; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 NYEMK 5
        |||||
Db      284 NYEMK 288

```

```

RESULT 15
US-60-465-241-4692
; Sequence 4692, Application US/60/465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-4692

```

```

Query Match          100.0%; Score 28; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 NYEMK 5
        |||||
Db      284 NYEMK 288

```

```

Search completed: July 11, 2003, 12:35:12
Job time : 5.79854 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 31.4761 Seconds

(without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-9

Perfect score: 50

Sequence: 1 AVAMMHQERK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*

8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*

9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*

10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*

11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*

12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*

13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*

14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*

15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*

16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*

17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*

18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*

19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*

20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*

21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*

22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*

23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*

24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*

25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*

26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*

27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	10	15	US-09-140-719-9
2	50	100.0	10	24	US-10-091-442-9
3	50	100.0	380	4	US-08-091-028A-34
4	50	100.0	380	15	US-09-140-719-34
5	50	100.0	380	19	US-09-508-997A-2
6	50	100.0	380	21	US-09-791-537-281

7	50	100.0	380	23	US-09-936-883A-2	Sequence 2, Appl1
8 <td>50 <td>100.0 <td>380 <td>24 <td>US-10-091-442-34</td> <td>Sequence 34, Appl1</td> </td></td></td></td>	50 <td>100.0 <td>380 <td>24 <td>US-10-091-442-34</td> <td>Sequence 34, Appl1</td> </td></td></td>	100.0 <td>380 <td>24 <td>US-10-091-442-34</td> <td>Sequence 34, Appl1</td> </td></td>	380 <td>24 <td>US-10-091-442-34</td> <td>Sequence 34, Appl1</td> </td>	24 <td>US-10-091-442-34</td> <td>Sequence 34, Appl1</td>	US-10-091-442-34	Sequence 34, Appl1
9 <td>50 <td>100.0 <td>380 <td>25 <td>US-10-110-030-2</td> <td>Sequence 2, Appl1</td> </td></td></td></td>	50 <td>100.0 <td>380 <td>25 <td>US-10-110-030-2</td> <td>Sequence 2, Appl1</td> </td></td></td>	100.0 <td>380 <td>25 <td>US-10-110-030-2</td> <td>Sequence 2, Appl1</td> </td></td>	380 <td>25 <td>US-10-110-030-2</td> <td>Sequence 2, Appl1</td> </td>	25 <td>US-10-110-030-2</td> <td>Sequence 2, Appl1</td>	US-10-110-030-2	Sequence 2, Appl1
10 <td>44 <td>88.0 <td>10</td> <td>4</td> <td>US-08-091-028-9</td> <td>Sequence 9, Appl1</td> </td></td>	44 <td>88.0 <td>10</td> <td>4</td> <td>US-08-091-028-9</td> <td>Sequence 9, Appl1</td> </td>	88.0 <td>10</td> <td>4</td> <td>US-08-091-028-9</td> <td>Sequence 9, Appl1</td>	10	4	US-08-091-028-9	Sequence 9, Appl1
11 <td>44 <td>88.0 <td>10</td> <td>4</td> <td>US-08-091-028A-9</td> <td>Sequence 4, Appl1</td> </td></td>	44 <td>88.0 <td>10</td> <td>4</td> <td>US-08-091-028A-9</td> <td>Sequence 4, Appl1</td> </td>	88.0 <td>10</td> <td>4</td> <td>US-08-091-028A-9</td> <td>Sequence 4, Appl1</td>	10	4	US-08-091-028A-9	Sequence 4, Appl1
12 <td>41 <td>82.0 <td>380</td> <td>19</td> <td>US-09-508-997A-4</td> <td>Sequence 19, Appl1</td> </td></td>	41 <td>82.0 <td>380</td> <td>19</td> <td>US-09-508-997A-4</td> <td>Sequence 19, Appl1</td> </td>	82.0 <td>380</td> <td>19</td> <td>US-09-508-997A-4</td> <td>Sequence 19, Appl1</td>	380	19	US-09-508-997A-4	Sequence 19, Appl1
13 <td>41 <td>82.0 <td>380</td> <td>23</td> <td>US-09-936-883A-19</td> <td>Sequence 6, Appl1</td> </td></td>	41 <td>82.0 <td>380</td> <td>23</td> <td>US-09-936-883A-19</td> <td>Sequence 6, Appl1</td> </td>	82.0 <td>380</td> <td>23</td> <td>US-09-936-883A-19</td> <td>Sequence 6, Appl1</td>	380	23	US-09-936-883A-19	Sequence 6, Appl1
14 <td>41 <td>82.0 <td>380</td> <td>25</td> <td>US-10-110-030-4</td> <td>Sequence 21, Appl1</td> </td></td>	41 <td>82.0 <td>380</td> <td>25</td> <td>US-10-110-030-4</td> <td>Sequence 21, Appl1</td> </td>	82.0 <td>380</td> <td>25</td> <td>US-10-110-030-4</td> <td>Sequence 21, Appl1</td>	380	25	US-10-110-030-4	Sequence 21, Appl1
15 <td>37</td> <td>74.0 <td>368</td> <td>19</td> <td>US-09-508-997A-6</td> <td>Sequence 21, Appl1</td> </td>	37	74.0 <td>368</td> <td>19</td> <td>US-09-508-997A-6</td> <td>Sequence 21, Appl1</td>	368	19	US-09-508-997A-6	Sequence 21, Appl1
16 <td>37</td> <td>74.0 <td>368</td> <td>25</td> <td>US-09-936-883A-21</td> <td>Sequence 21, Appl1</td> </td>	37	74.0 <td>368</td> <td>25</td> <td>US-09-936-883A-21</td> <td>Sequence 21, Appl1</td>	368	25	US-09-936-883A-21	Sequence 21, Appl1
17 <td>37</td> <td>74.0 <td>368</td> <td>25</td> <td>US-10-110-030-6</td> <td>Sequence 41683, A</td> </td>	37	74.0 <td>368</td> <td>25</td> <td>US-10-110-030-6</td> <td>Sequence 41683, A</td>	368	25	US-10-110-030-6	Sequence 41683, A
18 <td>36</td> <td>72.0 <td>118</td> <td>1</td> <td>PCT-US01-08631-41683</td> <td>Sequence 5860, Ap</td> </td>	36	72.0 <td>118</td> <td>1</td> <td>PCT-US01-08631-41683</td> <td>Sequence 5860, Ap</td>	118	1	PCT-US01-08631-41683	Sequence 5860, Ap
19 <td>36</td> <td>72.0 <td>336</td> <td>19</td> <td>US-09-540-209B-5860</td> <td>Sequence 16208, A</td> </td>	36	72.0 <td>336</td> <td>19</td> <td>US-09-540-209B-5860</td> <td>Sequence 16208, A</td>	336	19	US-09-540-209B-5860	Sequence 16208, A
20 <td>35</td> <td>70.0 <td>212</td> <td>21</td> <td>US-09-791-537-79944</td> <td>Sequence 2673, Ap</td> </td>	35	70.0 <td>212</td> <td>21</td> <td>US-09-791-537-79944</td> <td>Sequence 2673, Ap</td>	212	21	US-09-791-537-79944	Sequence 2673, Ap
21 <td>35</td> <td>70.0 <td>576</td> <td>21</td> <td>US-09-708-427-16208</td> <td>Sequence 1280, Ap</td> </td>	35	70.0 <td>576</td> <td>21</td> <td>US-09-708-427-16208</td> <td>Sequence 1280, Ap</td>	576	21	US-09-708-427-16208	Sequence 1280, Ap
22 <td>35</td> <td>70.0 <td>578</td> <td>21</td> <td>US-09-708-427-2673</td> <td>Sequence 328, App</td> </td>	35	70.0 <td>578</td> <td>21</td> <td>US-09-708-427-2673</td> <td>Sequence 328, App</td>	578	21	US-09-708-427-2673	Sequence 328, App
23 <td>35</td> <td>70.0 <td>602</td> <td>21</td> <td>US-09-708-427-16207</td> <td>Sequence 3046, Ap</td> </td>	35	70.0 <td>602</td> <td>21</td> <td>US-09-708-427-16207</td> <td>Sequence 3046, Ap</td>	602	21	US-09-708-427-16207	Sequence 3046, Ap
24 <td>35</td> <td>70.0 <td>1373</td> <td>27</td> <td>US-60-243-468-1280</td> <td>Sequence 19011, A</td> </td>	35	70.0 <td>1373</td> <td>27</td> <td>US-60-243-468-1280</td> <td>Sequence 19011, A</td>	1373	27	US-60-243-468-1280	Sequence 19011, A
25 <td>34</td> <td>68.0 <td>67</td> <td>13</td> <td>US-08-904-468-328</td> <td>Sequence 380, App</td> </td>	34	68.0 <td>67</td> <td>13</td> <td>US-08-904-468-328</td> <td>Sequence 380, App</td>	67	13	US-08-904-468-328	Sequence 380, App
26 <td>34</td> <td>68.0 <td>67</td> <td>13</td> <td>US-09-547-599C-3046</td> <td>Sequence 25985, A</td> </td>	34	68.0 <td>67</td> <td>13</td> <td>US-09-547-599C-3046</td> <td>Sequence 25985, A</td>	67	13	US-09-547-599C-3046	Sequence 25985, A
27 <td>34</td> <td>68.0 <td>76</td> <td>25</td> <td>US-10-155-881-19011</td> <td>Sequence 160, App</td> </td>	34	68.0 <td>76</td> <td>25</td> <td>US-10-155-881-19011</td> <td>Sequence 160, App</td>	76	25	US-10-155-881-19011	Sequence 160, App
28 <td>34</td> <td>68.0 <td>82</td> <td>16</td> <td>US-09-215-435-380</td> <td>Sequence 160, App</td> </td>	34	68.0 <td>82</td> <td>16</td> <td>US-09-215-435-380</td> <td>Sequence 160, App</td>	82	16	US-09-215-435-380	Sequence 160, App
29 <td>34</td> <td>68.0 <td>82</td> <td>22</td> <td>US-09-834-366-25985</td> <td>Sequence 41682, A</td> </td>	34	68.0 <td>82</td> <td>22</td> <td>US-09-834-366-25985</td> <td>Sequence 41682, A</td>	82	22	US-09-834-366-25985	Sequence 41682, A
30 <td>34</td> <td>68.0 <td>82</td> <td>22</td> <td>US-09-856-231A-123</td> <td>Sequence 3510, Ap</td> </td>	34	68.0 <td>82</td> <td>22</td> <td>US-09-856-231A-123</td> <td>Sequence 3510, Ap</td>	82	22	US-09-856-231A-123	Sequence 3510, Ap
31 <td>34</td> <td>68.0 <td>82</td> <td>27</td> <td>US-60-069-957-248</td> <td>Sequence 2881, Ap</td> </td>	34	68.0 <td>82</td> <td>27</td> <td>US-60-069-957-248</td> <td>Sequence 2881, Ap</td>	82	27	US-60-069-957-248	Sequence 2881, Ap
32 <td>34</td> <td>68.0 <td>82</td> <td>27</td> <td>US-60-177-646-2881</td> <td>Sequence 573, App</td> </td>	34	68.0 <td>82</td> <td>27</td> <td>US-60-177-646-2881</td> <td>Sequence 573, App</td>	82	27	US-60-177-646-2881	Sequence 573, App
33 <td>34</td> <td>68.0 <td>83</td> <td>1</td> <td>PCT-US98-12125-155</td> <td>Sequence 509, App</td> </td>	34	68.0 <td>83</td> <td>1</td> <td>PCT-US98-12125-155</td> <td>Sequence 509, App</td>	83	1	PCT-US98-12125-155	Sequence 509, App
34 <td>34</td> <td>68.0 <td>83</td> <td>16</td> <td>US-09-209-462B-160</td> <td>Sequence 44885, A</td> </td>	34	68.0 <td>83</td> <td>16</td> <td>US-09-209-462B-160</td> <td>Sequence 44885, A</td>	83	16	US-09-209-462B-160	Sequence 44885, A
35 <td>34</td> <td>68.0 <td>221</td> <td>1</td> <td>PCT-US01-08631-41682</td> <td>Sequence 85844, A</td> </td>	34	68.0 <td>221</td> <td>1</td> <td>PCT-US01-08631-41682</td> <td>Sequence 85844, A</td>	221	1	PCT-US01-08631-41682	Sequence 85844, A
36 <td>34</td> <td>68.0 <td>83</td> <td>26</td> <td>US-10-219-793-160</td> <td>Sequence 28004, A</td> </td>	34	68.0 <td>83</td> <td>26</td> <td>US-10-219-793-160</td> <td>Sequence 28004, A</td>	83	26	US-10-219-793-160	Sequence 28004, A
37 <td>34</td> <td>68.0 <td>980</td> <td>16</td> <td>US-09-252-991A-30838</td> <td></td> </td>	34	68.0 <td>980</td> <td>16</td> <td>US-09-252-991A-30838</td> <td></td>	980	16	US-09-252-991A-30838	
38 <td>33</td> <td>66.0 <td>70</td> <td>27</td> <td>US-60-177-571-3510</td> <td></td> </td>	33	66.0 <td>70</td> <td>27</td> <td>US-60-177-571-3510</td> <td></td>	70	27	US-60-177-571-3510	
39 <td>33</td> <td>66.0 <td>70</td> <td>27</td> <td>US-60-177-646-2881</td> <td></td> </td>	33	66.0 <td>70</td> <td>27</td> <td>US-60-177-646-2881</td> <td></td>	70	27	US-60-177-646-2881	
40 <td>33</td> <td>66.0 <td>70</td> <td>27</td> <td>US-60-178-311-575</td> <td></td> </td>	33	66.0 <td>70</td> <td>27</td> <td>US-60-178-311-575</td> <td></td>	70	27	US-60-178-311-575	
41 <td>33</td> <td>66.0 <td>108</td> <td>27</td> <td>US-60-140-803-509</td> <td></td> </td>	33	66.0 <td>108</td> <td>27</td> <td>US-60-140-803-509</td> <td></td>	108	27	US-60-140-803-509	
42 <td>33</td> <td>66.0 <td>161</td> <td>27</td> <td>US-60-142-845-609</td> <td></td> </td>	33	66.0 <td>161</td> <td>27</td> <td>US-60-142-845-609</td> <td></td>	161	27	US-60-142-845-609	
43 <td>33</td> <td>66.0 <td>275</td> <td>21</td> <td>US-09-791-537-44885</td> <td></td> </td>	33	66.0 <td>275</td> <td>21</td> <td>US-09-791-537-44885</td> <td></td>	275	21	US-09-791-537-44885	
44 <td>33</td> <td>66.0 <td>302</td> <td>21</td> <td>US-09-791-537-85844</td> <td></td> </td>	33	66.0 <td>302</td> <td>21</td> <td>US-09-791-537-85844</td> <td></td>	302	21	US-09-791-537-85844	
45 <td>33</td> <td>66.0 <td>303</td> <td>21</td> <td>US-09-791-537-28004</td> <td></td> </td>	33	66.0 <td>303</td> <td>21</td> <td>US-09-791-537-28004</td> <td></td>	303	21	US-09-791-537-28004	

ALIGNMENTS

RESULT 1

US-09-140-719-9

Sequence 9, Application US/09140719

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masaaki

APPLICANT: IMASA, Fuyuki

APPLICANT: TSUBOYOKA, Nobuo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, Nobuhiko

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAGUCHI, Kozo

APPLICANT: YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILED DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-9

Query Match 100.0%; Score 50; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
Db 1 AVAMMHQERK 10

RESULT 2

US-10-091-442-9

Sequence 9, Application US/10091442

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

TSUROOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhiko

KURIHARA, Tatsuya

YAMAICHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-091-442-9

Query Match 100.0%; Score 50; DB 24; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
Db 1 AVAMMHQERK 10

RESULT 3

US-08-091-028A-34

Sequence 34, Application US/08091028A

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

TSUROOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobuhiko

KURIHARA, Tatsuya

YAMAICHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: George Mason Bldg., Washington & Prince Sts.

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/091,028A

FILING DATE: 14-JUL-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: REA, TERESA STANEK

REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-091-028A-34

Query Match 100.0%; Score 50; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. NO. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMHDERK 10
DB 204 AVAMHDERK 213

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
GENERAL INFORMATION:
APPLICANT: TSUTSUMOTO, Masaaki
APPLICANT: IWASHI, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenji
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 50; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. NO. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMHDERK 10
DB 204 AVAMHDERK 213

RESULT 5
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2

Query Match 100.0%; Score 50; DB 19; Length 380;
Best Local Similarity 100.0%; Pred. NO. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMHDERK 10
DB 204 AVAMHDERK 213

RESULT 6
US-09-791-537-281
Sequence 281, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biocomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent version 3.0
SEQ ID NO 281
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-281

Query Match 100.0%; Score 50; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. NO. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMHDERK 10
DB 204 AVAMHDERK 213

RESULT 7
US-09-936-883A-2
; Sequence 2, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DPICT
; CURRENT APPLICATION NUMBER: US/09/936, 883A
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-936-883A-2

Query Match 100.0%; Score 50; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.21; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 AVAMHQRK 10
DB 204 AVAMHQRK 213

RESULT 8
US-10-091-442-34
; Sequence 34, Application US/10091442
; GENERAL INFORMATION:
; APPLICANT: TSUTIMOTO, Masafumi
; IMASA, Fuyuki
; TSURUOKA, NODUO
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, Nobuhiko
; KURIHARA, Tatsuya
; YAMAICHI, Kozo
; YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/091,442
; FILING DATE: 07-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/140,719
; FILING DATE: 08-AUG-1998
; APPLICATION NUMBER: US 08/474,661
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-247
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 50; DB 24; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.21; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 AVAMHQRK 10
DB 204 AVAMHQRK 213

RESULT 9
US-10-110-030-2
; Sequence 2, Application US/10110030
; GENERAL INFORMATION:
; APPLICANT: Miyata, Toshio
; TITLE OF INVENTION: ANIMAL MODEL FOR MESSENGIAL PROLIFERATIVE
; FILE REFERENCE: SHIM015
; CURRENT APPLICATION NUMBER: US/10/110,030
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 11/285736
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-110-030-2

Query Match 100.0%; Score 50; DB 25; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.21; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 1 AVAMHQRK 10
DB 204 AVAMHQRK 213

RESULT 10
US-08-091-028-9
; Sequence 9, Application US/08091028
; GENERAL INFORMATION:
; APPLICANT: TSUTIMOTO, Masafumi
; IMASA, Fuyuki
; TSURUOKA, NODUO
; NAKAZATO, Hiroshi
; MIURA, Kenju
; ISHIDA, Nobuhiko
; KURIHARA, Tatsuya
; YAMAICHI, Kozo
; YAMAGUCHI, Nozomi
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria

STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028
FILING DATE: 14-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA: JP 5-067339
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028-9

Query Match 88.0%; Score 44; DB 4; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.048;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVAMHQRK 10
Db 1 AVAMHQRK 10

RESULT 11
US-08-091-028A-9
Sequence 9, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSUBOOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhito
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A
FILING DATE: 14-JUL-1993

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA: JP 5-067339
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA SPANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-091-028A-9

Query Match 88.0%; Score 44; DB 4; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.048;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVAMHQRK 10
Db 1 AVAMHQRK 10

RESULT 12
US-09-508-997A-4
Sequence 4, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 380
TYPE: PRT
ORGANISM: Rattus rattus
FEATURE:
NAME/KEY: UNSURE
LOCATION: 51, 94
OTHER INFORMATION: xaa is unknown.
US-09-508-997A-4

Query Match 82.0%; Score 41; DB 19; Length 380;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVAMHQRK 10
Db 204 AVAMHQRK 213

RESULT 13
US-09-936-883A-19
Sequence 19, Application US/09936883A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
FILE REFERENCE: F2-101DIPCT
CURRENT APPLICATION NUMBER: US/09/936,883A

```

; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; PRIOR FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 51, 94
; OTHER INFORMATION: Xaa is unknown.
US-09-936-883A-19

```

```

Query Match      82.0%; Score 41; DB 23; Length 380;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AVAMMHQERK 10
        |||||||
DB      204 AVNMMHOERR 213

```

```

RESULT 14
US-10-110-030-4
; Sequence 4, Application US/10110030
; GENERAL INFORMATION:
; APPLICANT: Miyata, Toshio
; TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
; FILE REFERENCE: SHIM015
; CURRENT APPLICATION NUMBER: US/10/110,030
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 11/285736
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 51, 94
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-110-030-4

```

```

Query Match      82.0%; Score 41; DB 25; Length 380;
Best Local Similarity 80.0%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 AVAMMHQERK 10
        |||||||
DB      204 AVNMMHOERR 213

```

```

RESULT 15
US-09-508-997A-6
; Sequence 6, Application US/09508997A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: Mesgin Protein
; FILE REFERENCE: SHIM-004
; CURRENT APPLICATION NUMBER: US/09/508,997A
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: JP 09-275302
; PRIOR FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6

```

```

; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-508-997A-6

```

```

Query Match      74.0%; Score 37; DB 19; Length 368;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 VAMMHQERK 10
        |||||||
DB      193 VNMHHQERR 201

```

```

Search completed: July 11, 2003, 12:27:40
Job time : 33.4761 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 9.39709 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-9
Perfect score: 50
Sequence: 1 AVAMMHQERK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1229476 seqs, 263775074 residues

Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep4:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep4:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep4:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep4:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep4:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep4:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	50	100.0	380	10	US-09-936-883C-2
2	50	100.0	380	12	US-10-126-052A-263
3	50	100.0	380	14	US-60-443-566-3006
4	50	100.0	380	14	US-60-455-444-4692
5	50	100.0	380	14	US-60-465-241-4692
6	41	82.0	380	10	US-09-936-883C-19
7	40	80.0	131	10	US-09-864-408A-6424
8	37	74.0	368	10	US-09-936-883C-21
9	34	68.0	48	14	US-60-443-566-4420
10	34	68.0	69	14	US-60-443-566-4419
11	34	68.0	82	10	US-09-724-676-86949
12	34	68.0	82	10	US-60-443-566-4418
13	34	68.0	83	12	US-10-219-051B-5879
14	34	68.0	83	12	US-10-219-051B-5879
15	34	68.0	83	12	US-10-219-051B-5882
16	34	68.0	698	12	US-10-437-963-172615
17	34	68.0	980	12	US-10-366-683-30838
18	34	68.0	980	12	US-10-419-128-30838
19	33	66.0	161	12	US-10-437-963-160324

20	33	66.0	337	12	US-10-289-762-411	Sequence 411, App
21	33	66.0	453	12	US-10-437-963-104627	Sequence 104627, A
22	33	66.0	466	12	US-10-366-683-17916	Sequence 17916, A
23	33	66.0	466	12	US-10-419-128-17916	Sequence 17916, A
24	33	66.0	560	10	US-09-949-016-10786	Sequence 10786, A
25	33	66.0	560	10	US-10-087-192-1020	Sequence 22017, Ap
26	33	66.0	646	12	US-10-227-425-2207	Sequence 22017, Ap
27	33	66.0	655	14	US-60-452-680-24115	Sequence 24115, A
28	33	66.0	700	12	US-10-276-781-1352	Sequence 1352, Ap
29	33	66.0	700	12	US-10-366-547-63	Sequence 63, Appl
30	33	66.0	796	12	US-10-219-051B-1638	Sequence 12638, A
31	32	64.0	96	12	US-10-437-963-179299	Sequence 179299, A
32	32	64.0	136	12	US-10-264-237-2376	Sequence 2376, Ap
33	32	64.0	149	12	US-10-437-963-154295	Sequence 154295, A
34	32	64.0	220	12	US-10-424-599-210174	Sequence 210174, A
35	32	64.0	221	12	US-10-425-114-38163	Sequence 38163, A
36	32	64.0	235	12	US-10-266-887-19	Sequence 19, Appl
37	32	64.0	279	10	US-09-858-332A-14	Sequence 14, Appl
38	32	64.0	279	10	US-09-858-332B-14	Sequence 14, Appl
39	32	64.0	279	10	US-09-858-332C-14	Sequence 14, Appl
40	32	64.0	298	12	US-10-282-122A-62871	Sequence 62871, A
41	32	64.0	298	12	US-10-282-122A-64333	Sequence 64333, A
42	32	64.0	304	12	US-10-282-122A-62036	Sequence 62036, A
43	32	64.0	493	12	US-10-366-683-28024	Sequence 28024, A
44	32	64.0	493	12	US-10-419-128-28024	Sequence 28024, A
45	32	64.0	496	12	US-10-282-122A-76573	Sequence 76573, A

ALIGNMENTS

RESULT 1
US-09-936-883C-2
Sequence 2, Appl
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
FILE REFERENCE: P2-101DIPCT
CURRENT APPLICATION NUMBER: US/09/936, 883C
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883C-2

Query Match 100.0%; Score 50; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AVAMMHQERK 10
|||||||
Db 204 AVAMMHQERK 213

RESULT 2
US-10-126-052A-263
Sequence 263, Appl
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Murray, Richard
APPLICANT: Bos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
FILE REFERENCE: 018501-001530US
CURRENT APPLICATION NUMBER: US/10/126, 052A

```

1 CURRENT FILING DATE: 2002-04-18
2 -PRIORITY APPLICATION NUMBER: US 60/284,770
3 -PRIORITY FILING DATE: 2001-04-18
4 -PRIORITY APPLICATION NUMBER: US 60/290,452
5 -PRIORITY FILING DATE: 2001-05-10
6 -PRIORITY APPLICATION NUMBER: US 60/339,245
7 -PRIORITY FILING DATE: 2001-11-09
8 -PRIORITY APPLICATION NUMBER: US 60/350,666
9 -PRIORITY FILING DATE: 2001-11-13
10 -PRIORITY APPLICATION NUMBER: US 60/334,370
11 -PRIORITY FILING DATE: 2001-11-29
12 -PRIORITY APPLICATION NUMBER: US 60/372,246
13 -PRIORITY FILING DATE: 2002-04-12
14 -NUMBER OF SEQ ID NOS: 691
15 -SOFTWARE: PatentIn Ver. 2.1
16 SEQ ID NO 263
17 LENGTH: 380
18 TYPE: PRF
19 ORGANISM: Homo sapiens
20 US-10-126-052A-263

```

Query Match	100.0%;	Score 50;	DB 12;	Length 380;
Best Local Similarity	100.0%;	Pred. No. 0.12;		
Matches	10;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

```
QY      1 AVAMMHQERK 10
          |||||
Db      204 AVAMMHQERK 213
```

RESULT 3
US-60-443-566-3006
; Sequence 3006, Application US/60443566
; CENTRAL INFORMATION.

```

: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001447
: CURRENT APPLICATION NUMBER: US/60/443,566
: CURRENT FILING DATE: 2003-01-30
: NUMBER OF SEQ ID NOS: 25102
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3006
: LENGTH: 380
: TYPE: PR1
: ORGANISM: Homo sapiens
: US-60-443-566-3006

```

Query Match	100.0%	Score 50;	DB 14;	Length 380;
Best Local Similarity	100.0%;	Pred. No. 0 12;		
Matches	10;	Conservative 0;	Mismatches 0;	Indels 0;
QY	1	AVAMMHQERK	10	
Db	204	AVAMMHQERK	213	

Db . . 204 AVAMMHQERK 213

```

: RESULT 4
: US-60-455-444-4692
: Sequence 4692, Application US/60455444
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: BEGOVICH, Ann
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001455
: CURRENT APPLICATION NUMBER: US/60/455,444
: CURRENT FILING DATE: 2003-03-18
: NUMBER OF SEQ ID NOS: 50986
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4692
:
: LENGTH: 380

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-4692

```

Query Match	100.0%;	Score 50;	DB 14;	Length 380;
Best Local Similarity	100.0%;	Pred. NO. 0.12;		
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 AVAMMHQERK 10
          |||||
Db      204 AVAMMHQERK 213
```

RESULT 5
US-60-465-241-4692
; Sequence 4692, Application US/60465241
; Inventor: TROUBERT, ROY

```

: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001468
: CURRENT APPLICATION NUMBER: US/60/465,241
: CURRENT FILING DATE: 2003-04-23
: NUMBER OF SEQ ID NOS: 258418
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4692
: LENGTH: 380
: TYPE: PRF
: ORGANISM: Homo sapiens
: US-60-465-241-4692

```

ORGANISM: Homo sapiens
; US-60-465-241-4692

Query Match	100.0%;	Score 50;	DB 14;	Length 380;
Best Local Similarity	100.0%;	Pred. No. 0.12;		
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 AVAMMHQERK 10
          |||||
Db      204 AVAMMHQERK 213
```

RESULT 6
US-09-936-883C-19
; Sequence 19, Application US/09936883C
ORIGINAL TRANSMISSION

```

1  APPLICANT: MITTANI, JOSHUA
2  TITLE OF INVENTION: A Method for Detecting Meslin Protein and Use
3  TITLE OF INVENTION: Theoreof
4  FILE REFERENCE: P2-101DP1PCT
5  CURRENT APPLICATION NUMBER: US/09/936, 883C
6  CURRENT FILING DATE: 2001-12-21
7  PRIOR APPLICATION NUMBER: JP 1999-75305
8  PRIOR FILING DATE: 1999-03-19
9  PRIOR APPLICATION NUMBER: JP 1999-306623
10 PRIOR FILING DATE: 1999-10-28
11 NUMBER OF SEQ. ID NOS: 21
12 SOFTWARE: PatentIn Ver. 2.0
13 SEQ ID NO 19
14 LENGTH: 380

```

```
; SEQ ID NO 19
; LENGTH: 380
```

```

; ORGANISM: Rattus norvegicus
;
; FEATURE:
;
; NAME/KEY: misc_feature
;
; LOCATION: 51, 94
;
; OTHER INFORMATION: Xaa is u
US-09-936-883C-19

```

OTHER INFORMATION: Xaa is unknown.
US-09-936-883C-19

Query Match	82.0%;	Score 41;	DB 10;	Length 380;
Best Local Similarity	80.0%;	Pred. No. 6.9;		
Matches	8; Conservative	1; Mismatches	1; Indels	0; Gaps

QY 1 AVAMMHQERK 10
|| || || || || :

Db 204 AVMMHOERR 213

RESULT 7

US-09-864-408A-6424
; Sequence 6424, Application US/09864408A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: Novel Human Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6424
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-6424

Query Match 80.0%; Score 40; DB 10; Length 131;
Best Local Similarity 88.9%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AVAMHOER 9
111111111
Db 90 AVAMHOER 98

RESULT 8

US-09-936-883C-21
; Sequence 21, Application US/09936883C
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
; FILE REFERENCE: F2-101DIPCT
; CURRENT APPLICATION NUMBER: US/09/936,883C
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-936-883C-21

Query Match 74.0%; Score 37; DB 10; Length 368;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VAMMOHER 10
111111111
Db 193 VMMHOERR 201

RESULT 9

US-60-443-566-4420
; Sequence 4420, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566

; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4420
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-4420

Query Match 68.0%; Score 34; DB 14; Length 48;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMHOER 9
111111111
Db 18 AIAVHEER 26

RESULT 10

US-60-443-566-4419
; Sequence 4419, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4419
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-4419

Query Match 68.0%; Score 34; DB 14; Length 69;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMHOER 9
111111111
Db 18 AIAVHEER 26

RESULT 11

US-09-724-676-86949
; Sequence 86949, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 86949
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-86949

Query Match 68.0%; Score 34; DB 10; Length 82;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 AVAMHOER 9
111111111
Db 18 AIAVHEER 26

RESULT 12

```

US-09-724-676A-86949
; Sequence 86949, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09724, 676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 86949
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-86949

```

```

Query Match      68.0%; Score 34; DB 10; Length 82;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AVAMMHQER 9
      1:1:1:1:1:1
Db      18 AIAVLHEER 26

```

```

RESULT 13
US-60-443-566-4418
; Sequence 4418, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4418
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-4418

```

```

Query Match      68.0%; Score 34; DB 14; Length 82;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AVAMMHQER 9
      1:1:1:1:1:1
Db      18 AIAVLHEER 26

```

```

RESULT 14
US-10-219-051B-5879
; Sequence 5879, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Counties
; CURRENT APPLICATION NUMBER: US/10/219, 051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 5879
; LENGTH: 83

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(83)
; OTHER INFORMATION: Xaa=Unknown amino acid
US-10-219-051B-5879

```

```

Query Match      68.0%; Score 34; DB 12; Length 83;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AVAMMHQER 9
      1:1:1:1:1:1
Db      18 AIAVLHEER 26

```

```

RESULT 15
US-10-219-051B-5882
; Sequence 5882, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Counties
; CURRENT APPLICATION NUMBER: US/10/219, 051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 5882
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(83)
; OTHER INFORMATION: Xaa=Unknown amino acid
US-10-219-051B-5882

```

```

Query Match      68.0%; Score 34; DB 12; Length 83;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 AVAMMHQER 9
      1:1:1:1:1:1
Db      18 AIAVLHEER 26

```

```

Search completed: July 11, 2003, 12:35:13
Job time : 10.4971 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:54:32 ; Search time 1196.09 Seconds

(Without alignments)
204.833 Million cell updates/sec

Title: US-10-091-442-34

Perfect score: 1959

Sequence: 1 MASLAANAEFCNLFREMD.....FLFVIRKDIILFSGKVS CP 380

Scoring table:

BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCYUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1959	100.0	380	4	US-08-091-028A-34
2	1959	100.0	380	15	US-09-140-719-34
3	1959	100.0	380	19	US-09-508-997A-2
4	1959	100.0	380	21	US-09-791-537-281
5	1959	100.0	380	23	US-09-936-883A-2
6	1959	100.0	380	24	US-10-091-442-34

7	1959	100.0	380	25	US-10-110-030-2	Sequence 2, Appl 1
8	1477	75.4	380	19 <td>US-09-508-997A-4</td> <td>Sequence 19, Appl 1</td>	US-09-508-997A-4	Sequence 19, Appl 1
9	1477	75.4	380	23 <td>US-09-936-883A-19</td> <td>Sequence 4, Appl 1</td>	US-09-936-883A-19	Sequence 4, Appl 1
10	1477	75.4	380	25 <td>US-10-110-030-4</td> <td>Sequence 6, Appl 1</td>	US-10-110-030-4	Sequence 6, Appl 1
11	1458	74.4	368	19 <td>US-09-508-997A-6</td> <td>Sequence 21, Appl 1</td>	US-09-508-997A-6	Sequence 21, Appl 1
12	1458	74.4	368	23 <td>US-09-936-883A-21</td> <td>Sequence 6, Appl 1</td>	US-09-936-883A-21	Sequence 6, Appl 1
13	1458	74.4	368	25 <td>US-10-110-030-6</td> <td>Sequence 946, App</td>	US-10-110-030-6	Sequence 946, App
14	796.5	40.7	379	16 <td>US-09-852-911-946</td> <td>Sequence 946, App</td>	US-09-852-911-946	Sequence 946, App
15	796.5	40.7	379	22 <td>US-09-895-814-946</td> <td>Sequence 946, App</td>	US-09-895-814-946	Sequence 946, App
16	796.5	40.7	379	24 <td>US-10-006-920-946</td> <td>Sequence 946, App</td>	US-10-006-920-946	Sequence 946, App
17	796.5	40.7	379	24 <td>US-10-012-896-946</td> <td>Sequence 946, App</td>	US-10-012-896-946	Sequence 946, App
18	796.5	40.7	379	25 <td>US-10-144-678A-946</td> <td>Sequence 2, Appl 1</td>	US-10-144-678A-946	Sequence 2, Appl 1
19	771.5	39.4	395	22 <td>US-09-823-187-2</td> <td>Sequence 59, Appl 1</td>	US-09-823-187-2	Sequence 59, Appl 1
20	771.5	39.4	394	1 <td>PCT-US01-13360A-59</td> <td>Sequence 82, Appl 1</td>	PCT-US01-13360A-59	Sequence 82, Appl 1
21	769	39.3	394	24 <td>US-10-023-634-82</td> <td>Sequence 12, Appl 1</td>	US-10-023-634-82	Sequence 12, Appl 1
22	763.5	39.0	405	27 <td>US-60-341-362-12</td> <td>Sequence 83, Appl 1</td>	US-60-341-362-12	Sequence 83, Appl 1
23	763.5	39.0	423	24 <td>US-10-023-634-83</td> <td>Sequence 2, Appl 1</td>	US-10-023-634-83	Sequence 2, Appl 1
24	763.5	39.0	423	24 <td>PCT-US02-07215-2</td> <td>Sequence 2, Appl 1</td>	PCT-US02-07215-2	Sequence 2, Appl 1
25	754.5	38.5	425	24 <td>US-10-024-427-2</td> <td>Sequence 2, Appl 1</td>	US-10-024-427-2	Sequence 2, Appl 1
26	754.5	38.5	425	24 <td>US-10-094-944-2</td> <td>Sequence 2, Appl 1</td>	US-10-094-944-2	Sequence 2, Appl 1
27	754.5	38.5	425	27 <td>US-60-341-362-4</td> <td>Sequence 12, Appl 1</td>	US-60-341-362-4	Sequence 12, Appl 1
28	752	38.4	390	1 <td>PCT-US02-07215-12</td> <td>Sequence 2, Appl 1</td>	PCT-US02-07215-12	Sequence 2, Appl 1
29	752	38.4	390	4 <td>US-08-099-259-2</td> <td>Sequence 2, Appl 1</td>	US-08-099-259-2	Sequence 2, Appl 1
30	752	38.4	390	9 <td>US-08-568-147A-2</td> <td>Sequence 9, Appl 1</td>	US-08-568-147A-2	Sequence 9, Appl 1
31	752	38.4	390	20 <td>US-09-653-464B-9</td> <td>Sequence 3386, A</td>	US-09-653-464B-9	Sequence 3386, A
32	752	38.4	390	21 <td>US-09-791-537-3386</td> <td>Sequence 76123, A</td>	US-09-791-537-3386	Sequence 76123, A
33	752	38.4	390	21 <td>US-09-791-537-76123</td> <td>Sequence 121576,</td>	US-09-791-537-76123	Sequence 121576,
34	752	38.4	390	21 <td>US-09-791-537-76123</td> <td>Sequence 29, Appl 1</td>	US-09-791-537-76123	Sequence 29, Appl 1
35	752	38.4	390	21 <td>US-09-823-187-29</td> <td>Sequence 30, Appl 1</td>	US-09-823-187-29	Sequence 30, Appl 1
36	752	38.4	390	22 <td>US-09-823-187-30</td> <td>Sequence 65, Appl 1</td>	US-09-823-187-30	Sequence 65, Appl 1
37	752	38.4	390	22 <td>US-10-037-417-65</td> <td>Sequence 12, Appl 1</td>	US-10-037-417-65	Sequence 12, Appl 1
38	752	38.4	390	24 <td>US-10-094-944-12</td> <td>Sequence 106, App</td>	US-10-094-944-12	Sequence 106, App
39	752	38.4	413	21 <td>US-09-760-492-106</td> <td>Sequence 2, Appl 1</td>	US-09-760-492-106	Sequence 2, Appl 1
40	752	38.4	413	25 <td>US-10-162-749-106</td> <td>Sequence 13, Appl 1</td>	US-10-162-749-106	Sequence 13, Appl 1
41	752	38.4	425	25 <td>US-10-113-113-2</td> <td>Sequence 10, Appl 1</td>	US-10-113-113-2	Sequence 10, Appl 1
42	750.5	38.3	390	1 <td>PCT-US02-07215-13</td> <td></td>	PCT-US02-07215-13	
43	748	38.2	390	20 <td>US-09-653-464B-10</td> <td></td>	US-09-653-464B-10	
44	748	38.2	390	20 <td></td> <td></td>		
45	748	38.2	390	20 <td></td> <td></td>		

ALIGNMENTS

RESULT 1
US-08-091-028A-34
Sequence 34, Application US/08091028A
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMWASA, Fuyuki
APPLICANT: TSURUOKA, Nobuo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, Nobuhiko
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAIUCHI, Kozo
APPLICANT: YAMAGUCHI, Nozomi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/091,028A

FILING DATE: 14-JUL-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 4-212305
 FILING DATE: 17-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 5-067339
 FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: REA, TERESA STANER
 REGISTRATION NUMBER: 30,427
 REFERENCE/DOCKET NUMBER: 001560-204
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-6620
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 380 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-091-028A-34

Query Match 100.0%; Score 1959; DB 4; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.1e-179;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 DB 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 QY 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHKDYIECAEK 120
 DB 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTNHLDTNRNINIKWVENETHGKIKVYIGGGISSAVMLVNAVYFKGK 180
 DB 121 LYDAKVERVDFTNHLDTNRNINIKWVENETHGKIKVYIGGGISSAVMLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFPKSCGKAVAMHOERKFNLSVIDPMSKILLETRYNGIMNYVL 240
 DB 181 WQSAFTKSETINCHFPKSCGKAVAMHOERKFNLSVIDPMSKILLETRYNGIMNYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDF 300
 DB 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDF 300
 QY 301 DESKADLSGIASGGRILYISRMHKSYTEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGGRILYISRMHKSYTEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSCP 380
 DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 2
 US-09-140-719-34
 Sequence 34, Application US/09140719
 GENERAL INFORMATION:
 APPLICANT: TSUJIMOTO, Masafumi
 APPLICANT: IWASA, Fuyuki
 APPLICANT: TSUROOKA, Nobuo
 APPLICANT: NAKAZATO, Hiroshi
 APPLICANT: MURA, Kenju
 APPLICANT: ISHIDA, Nobuhito
 APPLICANT: KURIHARA, Tatsuya
 APPLICANT: YAMACHI, Kozo
 APPLICANT: YAMAGUCHI, Nozomi
 TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/140,719
 FILING DATE: 08-AUG-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/474,661
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/091,028
 FILING DATE: 14-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 4-212305
 FILING DATE: 17-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-067339
 FILING DATE: 04-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 001560-247
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 380 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-140-719-34

Query Match 100.0%; Score 1959; DB 15; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.1e-179;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 DB 1 MASLAANAEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGAQDDSLSDIKLHV 60
 QY 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHKDYIECAEK 120
 DB 61 NTASGYGNSNSQSGLOSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHKDYIECAEK 120
 QY 121 LYDAKVERVDFTNHLDTNRNINIKWVENETHGKIKVYIGGGISSAVMLVNAVYFKGK 180
 DB 121 LYDAKVERVDFTNHLDTNRNINIKWVENETHGKIKVYIGGGISSAVMLVNAVYFKGK 180
 QY 181 WQSAFTKSETINCHFPKSCGKAVAMHOERKFNLSVIDPMSKILLETRYNGIMNYVL 240
 DB 181 WQSAFTKSETINCHFPKSCGKAVAMHOERKFNLSVIDPMSKILLETRYNGIMNYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDF 300
 DB 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDF 300
 QY 301 DESKADLSGIASGGRILYISRMHKSYTEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
 DB 301 DESKADLSGIASGGRILYISRMHKSYTEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSCP 380
 DB 361 FLFVIRKDDIILFSGKVSCP 380

RESULT 3
US-09-508-997A-2
Sequence 2, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508, 997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-997A-2

Query Match 100.0%; Score 1959; DB 19; Length 380;
Best Local Similarity 100.0%; Pred. No. 1,1e-179;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAEEFCFNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
DB 1 MASLAANAEEFCFNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
QY 61 NTASGYGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLFAEKYGGFHKDYIECAEK 120
DB 61 NTASGYGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLFAEKYGGFHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLLEPTRRINKKVENETHGKTKNYIGEGGSSAVMVLVNAVYFRKG 180
DB 121 LYDAKVERVDFTNLLEPTRRINKKVENETHGKTKNYIGEGGSSAVMVLVNAVYFRKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIEDPSKILELRNGGINNYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIEDPSKILELRNGGINNYVL 240
QY 241 LPENDLSEIENKLTFFQNLMEWTNPRMTSKYVEVFFPOFKIEKYEKKOYLRLALGLDIF 300
DB 241 LPENDLSEIENKLTFFQNLMEWTNPRMTSKYVEVFFPOFKIEKYEKKOYLRLALGLDIF 300
QY 301 DESADLSGASGRLYISRMHKSYLEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
DB 301 DESADLSGASGRLYISRMHKSYLEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
QY 361 FLVYIRKDDIILFSGKVSFCP 380
DB 361 FLVYIRKDDIILFSGKVSFCP 380

RESULT 4
US-09-791-537-281
Sequence 281, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomimix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent In Ver. 3.0
SEQ ID NO 281
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-281

Query Match 100.0%; Score 1959; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 1,1e-179;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAEEFCFNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
DB 1 MASLAANAEEFCFNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
QY 61 NTASGYGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLFAEKYGGFHKDYIECAEK 120
DB 61 NTASGYGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLFAEKYGGFHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLLEPTRRINKKVENETHGKTKNYIGEGGSSAVMVLVNAVYFRKG 180
DB 121 LYDAKVERVDFTNLLEPTRRINKKVENETHGKTKNYIGEGGSSAVMVLVNAVYFRKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIEDPSKILELRNGGINNYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIEDPSKILELRNGGINNYVL 240
QY 241 LPENDLSEIENKLTFFQNLMEWTNPRMTSKYVEVFFPOFKIEKYEKKOYLRLALGLDIF 300
DB 241 LPENDLSEIENKLTFFQNLMEWTNPRMTSKYVEVFFPOFKIEKYEKKOYLRLALGLDIF 300
QY 301 DESADLSGASGRLYISRMHKSYLEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
DB 301 DESADLSGASGRLYISRMHKSYLEVTEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
QY 361 FLVYIRKDDIILFSGKVSFCP 380
DB 361 FLVYIRKDDIILFSGKVSFCP 380

RESULT 5
US-09-936-883A-2
Sequence 2, Application US/09936883A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A method for Detecting Megsin Protein and Use
FILE REFERENCE: F2-101DPICT
CURRENT APPLICATION NUMBER: US/09/936, 883A
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883A-2
Query Match 100.0%; Score 1959; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 1,1e-179;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASLAANAEEFCFNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
DB 1 MASLAANAEEFCFNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLLHV 60
QY 61 NTASGYGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLFAEKYGGFHKDYIECAEK 120
DB 61 NTASGYGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLFAEKYGGFHKDYIECAEK 120
QY 121 LYDAKVERVDFTNLLEPTRRINKKVENETHGKTKNYIGEGGSSAVMVLVNAVYFRKG 180
DB 121 LYDAKVERVDFTNLLEPTRRINKKVENETHGKTKNYIGEGGSSAVMVLVNAVYFRKG 180
QY 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIEDPSKILELRNGGINNYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIEDPSKILELRNGGINNYVL 240

Db 181 WOSAFKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKLLELRNGGINMYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFFPOFKIEKNYEMKOYLALGKIDIF 300
 Db 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFFPOFKIEKNYEMKOYLALGKIDIF 300
 QY 301 DESKADLSGIASGGRLYISRMMHKSXYIEVTEEGTEATAAGSNIYERKOLPOSTLFRADHP 360
 Db 301 DESKADLSGIASGGRLYISRMMHKSXYIEVTEEGTEATAAGSNIYERKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSCEP 380
 Db 361 FLFVIRKDDIILFSGKVSCEP 380

RESULT 6

US-10-091-442-34

Sequence 34, Application US/10091442
 GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

IMASA, Fuyuki

TSUROJOKA, Nobuo

NAKAZATO, Hiroshi

MURA, Kenju

ISHIDA, Nobunhiro

KURIHARA, Tatsuya

YAMAICHI, Kozo

YAMAGUCHI, Nozomi

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-Aug-1998

APPLICATION NUMBER: US 08/474,661

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-247

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 34:

US-10-091-442-34

Query Match 100.0%; Score 1959; DB 24; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1,1e-179;
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREDDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
 Db 1 MASLAANAEEFCNLFREDDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
 QY 61 NTASGYGNSNSGSLQSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHNDYIECAEK 120
 Db 61 NTASGYGNSNSGSLQSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHNDYIECAEK 120
 QY 121 LYDAKVERVDFTNHLIEDTRRNINKWVENETHGKIKNVIGEGGISSAVVLYNAVYFKG 180
 Db 121 LYDAKVERVDFTNHLIEDTRRNINKWVENETHGKIKNVIGEGGISSAVVLYNAVYFKG 180
 QY 181 WOSAFKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKLLELRNGGINMYVL 240
 Db 181 WOSAFKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKLLELRNGGINMYVL 240
 QY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFFPOFKIEKNYEMKOYLALGKIDIF 300
 Db 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFFPOFKIEKNYEMKOYLALGKIDIF 300
 QY 301 DESKADLSGIASGGRLYISRMMHKSXYIEVTEEGTEATAAGSNIYERKOLPOSTLFRADHP 360
 Db 301 DESKADLSGIASGGRLYISRMMHKSXYIEVTEEGTEATAAGSNIYERKOLPOSTLFRADHP 360
 QY 361 FLFVIRKDDIILFSGKVSCEP 380
 Db 361 FLFVIRKDDIILFSGKVSCEP 380

RESULT 7

US-10-110-030-2

Sequence 2, Application US/10110030
 GENERAL INFORMATION:

APPLICANT: Miyata, Toshio

TITLE OF INVENTION: ANIMAL MODEL FOR MESENCHYMAL PROLIFERATIVE

FILE REFERENCE: SHIM015

CURRENT APPLICATION NUMBER: US/10/110,030

CURRENT FILING DATE: 2002-04-03

PRIOR APPLICATION NUMBER: 11/285736

PRIOR FILING DATE: 1999-10-06

NUMBER OF SEQ ID NOS: 15

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 2

LENGTH: 380

TYPE: PRT

ORGANISM: Homo sapiens

US-10-110-030-2

Query Match

Best Local Similarity

Matches 380; Conservative

100.0%; Score 1959; DB 25; Length 380;

Pred. No. 1,1e-179;

Mismatches 0; Indels 0; Gaps 0;

QY 1 MASLAANAEEFCNLFREDDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60

Db 1 MASLAANAEEFCNLFREDDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60

QY 61 NTASGYGNSNSGSLQSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHNDYIECAEK 120

Db 61 NTASGYGNSNSGSLQSLKRVFSDINASHKDYDLSYNGLFAEKVYGFHNDYIECAEK 120

QY 121 LYDAKVERVDFTNHLIEDTRRNINKWVENETHGKIKNVIGEGGISSAVVLYNAVYFKG 180

Db 121 LYDAKVERVDFTNHLIEDTRRNINKWVENETHGKIKNVIGEGGISSAVVLYNAVYFKG 180

QY 181 WOSAFKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKLLELRNGGINMYVL 240

Db 181 WOSAFKSETINCHFKSPKCSGKAVAMHQRKFNLSVIEDPSMKLLELRNGGINMYVL 240

OY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPPOFKTEKVENKQYLRALGLDIF 300
DB 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPPOFKTEKVENKQYLRALGLDIF 300
OY 301 DESRADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIVEKOLPOSTLEFRADHP 360
DB 301 DESRADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIVEKOLPOSTLEFRADHP 360
OY 361 FLFVIRKDDITLFGSKVSCP 380
DB 361 FLFVIRKDDITLFGSKVSCP 380

RESULT 8

US-09-508-997A-4
Sequence 4, Application US/09508997A
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: Megsin Protein
FILE REFERENCE: SHIM-004
CURRENT APPLICATION NUMBER: US/09/508,997A
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: JP 09-275302
PRIOR FILING DATE: 1997-09-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 380
TYPE: PRT
ORGANISM: Rattus rattus
FEATURE:
NAME/KEY: UNSURE
LOCATION: 51, 94
OTHER INFORMATION: Xaa is unknown.
US-09-508-997A-4

Query Match 75.4%; Score 1477; DB 19; Length 380;
Best Local Similarity 73.4%; Pred. No. 3,8e-133;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

OY 1 MASLAANAEEFCNLFRMDNNGNVEFSSLSLFAALAVRLGAODDSLQIDKLLHV 60
DB 1 MASLAANAEEFCNLFRMDNNGNVEFSSLSLFAALAVRLGAODDSLQIDKLLHV 60
OY 61 ISPSRQGNSSNSQGLQYQLKRVLADINSSHKDKLSTANGVPAEKYFDFHKSYMECHEN 120
DB 61 ISPSRQGNSSNSQGLQYQLKRVLADINSSHKDKLSTANGVPAEKYFDFHKSYMECHEN 120
OY 121 LYDAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
DB 121 LYDAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
OY 121 LYNAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
DB 121 LYNAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
OY 181 WSAFTYSETINCHFKSPKSCGKAVAMMHQERKNLSYIEDPSKLTLELRNGGINMYVL 240
DB 181 WSAFTYSETINCHFKSPKSCGKAVAMMHQERKNLSYIEDPSKLTLELRNGGINMYVL 240
OY 181 WSAFTYSDTLSCHFSPSPGKAVAMMHQERKNLSYIEDPSKLTLELRNGGINMYVL 240
DB 181 WSAFTYSDTLSCHFSPSPGKAVAMMHQERKNLSYIEDPSKLTLELRNGGINMYVL 240
OY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPPOFKTEKVENKQYLRALGLDIF 300
DB 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPPOFKTEKVENKQYLRALGLDIF 300
OY 301 DESRADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIVEKOLPOSTLEFRADHP 360
DB 301 DESRADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIVEKOLPOSTLEFRADHP 360
OY 361 FLFVIRKDDITLFGSKVSCP 380
DB 361 FLFVIRKDDITLFGSKVSCP 380

RESULT 9
US-09-936-883A-19
Sequence 19, Application US/09936883A
GENERAL INFORMATION:

APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
FILE REFERENCE: P2-101DPIPT
CURRENT APPLICATION NUMBER: US/09/936,883A
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 380
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc-feature
LOCATION: 51, 94
OTHER INFORMATION: Xaa is unknown.
US-09-936-883A-19

Query Match 75.4%; Score 1477; DB 23; Length 380;
Best Local Similarity 73.4%; Pred. No. 3,8e-133;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

OY 1 MASLAANAEEFCNLFRMDNNGNVEFSSLSLFAALAVRLGAODDSLQIDKLLHV 60
DB 1 MASLAANAEEFCNLFRMDNNGNVEFSSLSLFAALAVRLGAODDSLQIDKLLHV 60
OY 61 ISPSRQGNSSNSQGLQYQLKRVLADINSSHKDKLSTANGVPAEKYFDFHKSYMECHEN 120
DB 61 ISPSRQGNSSNSQGLQYQLKRVLADINSSHKDKLSTANGVPAEKYFDFHKSYMECHEN 120
OY 121 LYDAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
DB 121 LYDAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
OY 121 LYNAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
DB 121 LYNAKVERVDFTNHLIEDTRNINKVVENETHGKIKNVIGEGISSAAVMVLVNAVYFEGK 180
OY 181 WSAFTYSETINCHFKSPKSCGKAVAMMHQERKNLSYIEDPSKLTLELRNGGINMYVL 240
DB 181 WSAFTYSDTLSCHFSPSPGKAVAMMHQERKNLSYIEDPSKLTLELRNGGINMYVL 240
OY 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPPOFKTEKVENKQYLRALGLDIF 300
DB 241 LPENDLSEIENKLTFFONLMEWTNPRMTSKYVEVFPPOFKTEKVENKQYLRALGLDIF 300
OY 301 DESRADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIVEKOLPOSTLEFRADHP 360
DB 301 DESRADLSGIASGRLYISRMHKSYLEVTEEGTEATATGSNIVEKOLPOSTLEFRADHP 360
OY 361 FLFVIRKDDITLFGSKVSCP 380
DB 361 FLFVIRKDDITLFGSKVSCP 380

RESULT 10
US-10-110-030-4
Sequence 4, Application US/10110030
GENERAL INFORMATION:
APPLICANT: Miyata, Toshio
TITLE OF INVENTION: ANIMAL MODEL FOR MESENCHYMAL PROLIFERATIVE
FILE REFERENCE: SHIM015
CURRENT APPLICATION NUMBER: US/10/110,030
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: 11/285736
PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 380
TYPE: PRT
ORGANISM: Rattus norvegicus

FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 51, 94
; OTHER INFORMATION: Xaa - Any Amino Acid
US-10-110-030-4

Query Match 75.4%; Score 1477; DB 25; Length 380;
Best Local Similarity 73.4%; Pred. No. 3.8e-133;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

QY 1 MASLAANAFECNLTREMDNNGNNGVNFSSLSLFAALALVRLGAODDSLQIDKLHVMTASGNGSSNS 72
D 1 MASLAANAFECNLTREMDNNGNNGVNFSSLSLFAALALVRLGAODDSLQIDKLHVMTASGNGSSNS 72
QY 61 MPASGYGNSNSGSLQSLKRYFSDINASHKDYDLSTVNGLFPAEKYGFHKKDYIECAEK 120
D 61 ISPSRQGNSSNGLQIYQKRLADINSSHKDKSLANGVFAEKYFEDHKKYMECAEN 120
QY 121 LYDAKVERDFTNHLDETRRNKKNVENVETHGKIKVNGIGGSISSAVMLVNAVYFCKG 180
D 121 LYNAKVERDFTNHLDETRRNKKNVENVETHGKIKVNGIGGSISSAVMLVNAVYFCKG 180
QY 181 MOFAFKSETINCHPSPKSGKAVAMMOERKFNLSYEDPSMKITELRYNGGINMYVL 240
D 181 WKSAFTKSDTSLCHFPSPSGPKAVAMMOERKFNLSYEDPSMKITELRYNGGINMYVL 240
QY 241 LPENDLSELENKLTFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDIF 300
D 241 LPEDDLSELESKLSFQNLMDWTNSRKMKSOYVNFPLQFIEEDYERSHLSKVGLEDIF 300
QY 301 DESKADLSGIASGRLYISRMHKSLEYTEEGTEATATGNSIVVEKOLPSTLFRADHP 360
D 301 VESRADLSGIASGRLYISRMHKSLEYTEEGTEATATGNSIVVEKOLPSTLFRADHP 360
QY 361 FLFVIRKDDITLFSKVSCP 380
D 361 FLFVIRKDDITLFSKVSCP 380

RESULT 11.

US-09-508-997A-6
; Sequence 6, Application US/09508997A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: Megin Protein
; FILE REFERENCE: SHIM-004
; CURRENT APPLICATION NUMBER: US/09/508, 997A
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: JP 09-275302
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-508-997A-6

Query Match 74.4%; Score 1458; DB 19; Length 368;
Best Local Similarity 73.4%; Pred. No. 2.5e-131;
Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;

QY 13 FNLFRMDNNGNNGVNFSSLSLFAALALVRLGAODDSLQIDKLHVMTASGNGSSNS 72
D 1 FDLFRMDSSGNGNNGVNFSSLSLFAALALVRLGAODDSLQIDKLHVMTASGNGSSNS 72
QY 73 QSGLSQSLKRYFSDINASHKDYDLSTVNGLFPAEKYGFHKKDYIECAEKLYDAKVERDFT 132
D 61 QGQLQYOLKRLADINSSHKDYELSTATGVFAEKYDFHKNYTECAENLYNAKVERDFT 120
QY 133 NHEDETRRNKKNVENVETHGKIKVNGIGGSISSAVMLVNAVYFCKGQSAFTKSETIN 192
D 121 NOVQDTRFRKINKEIENETHGKIKVNGIGDSSLSAVMLVNAVYFCKGKWSAFTKDTLS 180

QY 193 CHFKSPKSGKAVAMMOERKFNLSYEDPSMKITELRYNGGINMYVLPLENDLSELENK 252
D 181 CHFRSPKCGKAVYVAMMOERKFNLSYEDPSMKITELRYNGGINMYVLPLENDLSELENK 240
QY 253 LFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDIFDESKADLSGIAS 312
D 241 LSFQNLMDWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDIFDESKADLSGIAS 300
QY 313 GGRLYISRMHKSLEYTEEGTEATATGNSIVVEKOLPSTLFRADHPFLYIRKDDITL 372
D 301 GGRLYISRMHKSLEYTEEGTEATATGNSIVVEKOLPSTLFRADHPFLYIRKDDITL 360
QY 373 FSGKVSCP 380
D 361 FTGKVSCP 368

RESULT 12

US-09-936-883A-21
; Sequence 21, Application US/09936883A
; GENERAL INFORMATION:
; APPLICANT: MIYATA, Toshio
; TITLE OF INVENTION: A Method for Detecting Megin Protein and Use
; FILE REFERENCE: F2-101DPICT
; CURRENT APPLICATION NUMBER: US/09/936, 883A
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: JP 1999-75305
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: JP 1999-306623
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-936-883A-21

Query Match 74.4%; Score 1458; DB 23; Length 368;
Best Local Similarity 73.4%; Pred. No. 2.5e-131;
Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;

QY 13 FNLFRMDNNGNNGVNFSSLSLFAALALVRLGAODDSLQIDKLHVMTASGNGSSNS 72
D 1 FDLFRMDSSGNGNNGVNFSSLSLFAALALVRLGAODDSLQIDKLHVMTASGNGSSNS 72
QY 73 QSGLSQSLKRYFSDINASHKDYDLSTVNGLFPAEKYGFHKKDYIECAEKLYDAKVERDFT 132
D 61 QGQLQYOLKRLADINSSHKDYELSTATGVFAEKYDFHKNYTECAENLYNAKVERDFT 120
QY 133 NHEDETRRNKKNVENVETHGKIKVNGIGGSISSAVMLVNAVYFCKGQSAFTKSETIN 192
D 121 NOVQDTRFRKINKEIENETHGKIKVNGIGDSSLSAVMLVNAVYFCKGKWSAFTKDTLS 180
QY 193 CHFKSPKSGKAVAMMOERKFNLSYEDPSMKITELRYNGGINMYVLPLENDLSELENK 252
D 181 CHFRSPKCGKAVYVAMMOERKFNLSYEDPSMKITELRYNGGINMYVLPLENDLSELENK 240
QY 253 LFFONLMEWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDIFDESKADLSGIAS 312
D 241 LSFQNLMDWTNPRMTSKYVEVFPOFKIEKNYEMKOYLALGLKIDIFDESKADLSGIAS 300
QY 313 GGRLYISRMHKSLEYTEEGTEATATGNSIVVEKOLPSTLFRADHPFLYIRKDDITL 372
D 301 GGRLYISRMHKSLEYTEEGTEATATGNSIVVEKOLPSTLFRADHPFLYIRKDDITL 360
QY 373 FSGKVSCP 380
D 361 FTGKVSCP 368

```

RESULT 13
US-10-110-030-6
: Sequence 6, Application US/10110030
: GENERAL INFORMATION:
: APPLICANT: Miyata, Toshio
: TITLE OF INVENTION: ANIMAL MODEL FOR MESENTERIAL PROLIFERATIVE
: FILE REFERENCE: SHIM015
: CURRENT APPLICATION NUMBER: US/10/110,030
: PRIOR FILING DATE: 1999-10-06
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 368
: TYPE: PRT
: ORGANISM: Mus musculus
US-10-110-030-6

```

Query Match 74.4% Score 1458 DB 25 Length 368

Best Local Similarity 73.4% Pred. No. 2.5e-131

Matches 270; Conservative 53; Mismatches 45; Indels 0; Gaps 0;

```

QY 13 FNLREMDNDGNGNNGVFFSSLSLFAALALVRLGAQDLSQIDKLHVTNAGSGNSNS 72
DB 1 FDLFREMDSQNGNNGVFFSSLSLFTALTRLRAGRCARQIDKALHFNIPRQGNSSNN 60
QY 73 QSLQSLQKRVFSDINASHKDYSLVNGLFAEKVGFHKDYIECAEKLYDAKVERVDT 132
DB 61 QPGLQYOLKRVLDINSHKDYELSLTGVFAEKVGFHKDYIECAENLYAKVERVDT 120
QY 133 NHELETRRNKNWENTHGRKIKVIGEGGISSSAVAVLVNAVYFKGKMSAFKSTPT 192
DB 121 NDVODTRFKINKNTENETHKIKKVLDSSSSAVAVLVNAVYFKGKMSAFKSTPTLS 180
QY 193 CHFKSPKSGKAAVAMHOKERKFNLSYEDPSMKILELRNGJINMYVLLPNDLSETEK 252
DB 181 CRFRSPKCPGKVVNMHGERFNLSIQPRMQLDQYNGISMYIMLPEDGCELESK 240
QY 253 LTFONLMEWTPRRMTSKYVEVFPPOKTEKNTKMOYLRALGLKIDFDESKADLSIAS 312
DB 241 LSFONLMDWTNRKMKSQYVNVLPQPKIEKNYEMTHLKLGLKIDFDESSADLSIAS 300
QY 313 GGRLYISRMHKSIVTEVTEGTEATATGSNIYEKOLPOSTFRADHPFLVIRKDDIIL 372
DB 301 GGRLYISKLMHKSIVTEGTEATATENNIVEKOLPESTVEPADPFLFLVIRKDDIIL 360
QY 373 FSGKVSQP 380
DB 361 FTGKVSQP 368

```

```

RESULT 14
US-09-283-946-2
: Sequence 2, Application US/09283946
: GENERAL INFORMATION:
: APPLICANT: Afar, Daniel E.
: APPLICANT: Hubert, Rene S.
: APPLICANT: Leong, Kahan
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Saffran, Douglas C.
: TITLE OF INVENTION: PROSTAPIN GENE AND PROTEIN AND USES THEREOF
: FILE REFERENCE: 1703-005.US1
: CURRENT APPLICATION NUMBER: US/09/283,946
: CURRENT FILING DATE: 1999-04-01
: EARLIER APPLICATION NUMBER: 60/080,167
: EARLIER FILING DATE: 1998-03-31
: EARLIER APPLICATION NUMBER: 60/085,720
: EARLIER FILING DATE: 1998-05-15
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn ver. 2.0
: SEQ ID NO 2

```

```

: LENGTH: 379
: TYPE: PRT
: ORGANISM: HUMAN PROSTAPIN CONSENSUS OF AMINO ACID SEQUENCE
US-09-283-946-2

```

Query Match 40.7% Score 796.5 DB 16 Length 379

Best Local Similarity 42.6% Pred. No. 2e-67

Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

```

QY 1 MASIAANAEFCENLREMDNDGNGVFFSSLSLFAALALVRLGAQDLSQIDKLHAY 60
DB 1 MGLSTANVEFCIDVFEKELNSNNGIDNIFESSLSLVALSMVLGAGETAEOLKVLAF 60
QY 61 -----NTASGYGNSN--SOSG-LOSOLKRVFSDINASHKDYSLVNGLFAEKVGFHKD 113
DB 61 SHYVDSLKPFKDSQKSGRHSERGVESQINODPSCTSIANRLIGTKTMAFHQQ 120
QY 114 YIDCAEKLYDAKVERVDTNHELETRRNKNWENTHGRKIKVIGBGISSAVAVLVN 173
DB 121 YLSCSEKMYQARLQTVDFEOSTEETRKMINAWENKTNGKAVANLFGKSTIDPSSVAVLVN 180
QY 174 AVYFKGWSQAFKSTETINCHFKSPKSGK--AVAMHOKERKFNLSYEDPSMKILELRY 231
DB 181 TIYFKGGRONKF-----QGRNVYEMVYQITFKLAFKVEKPOHQVLELPY 225
QY 232 -NGGINVYLLPEN--DLSEIENKLTFRONLMENTNPRMTSKYVEVFPPOKTEKNTK 288
DB 226 VNNKLSMIIILPVGIANLKOIEKOLNSGTHEWTSSNMHREVEVILPFEKLEIKYELN 265
QY 289 QYLRALGLKIDFDESKADLSIASGRLYISRMHKSIVTEVTEGTEATATGSNIYEKQ 348
DB 286 SLKPLGVTDLDEQVAKDLGMSPTKGLYSKAIHKSILYDVSEGTBAATGDSIAVKS 345
QY 349 LPPOSTFRADHPFLVIR--KDIILFSGVSCP 380
DB 346 LPMRAQFKANHPFLFLIRHTHTNITLFCGKLASP 379

```

```

RESULT 15
US-09-852-911-946
: Sequence 946, Application US/09852911
: GENERAL INFORMATION:
: APPLICANT: Xu, Jianshun
: APPLICANT: Dillon, David C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Jiang, Yugu
: APPLICANT: Henderson, Robert A.
: APPLICANT: Kalos, Michael D.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Retter, Marc W.
: APPLICANT: Stoik, John A.
: APPLICANT: Day, Craig H.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Carter, Darick
: APPLICANT: Li, Samuel
: APPLICANT: Wang, Aijun
: APPLICANT: Skeiky, Yasser A.W.
: APPLICANT: Hepler, William
: APPLICANT: Hurai, John
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Houghton, Raymond L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.427C25
: CURRENT APPLICATION NUMBER: US/09/852,911
: CURRENT FILING DATE: 2001-05-09
: NUMBER OF SEQ ID NOS: 947
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 946
: LENGTH: 379
: TYPE: PRT
: ORGANISM: Homo sapiens

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:55:27 ; Search time 357.089 Seconds
(without alignments)
280.699 Million cell updates/sec

Title: US-10-091-442-34
Perfect score: 1959
Sequence: 1 MASLAANAEEFCNLEFREM.....FLFVIRKDDILFSGKVSQP 380

Scoring table: BLOSUM62
Gap: 10.0, Gapext 0.5

Searched: 1229476 seqs, 263775074 residues
Total number of hits satisfying chosen parameters: 1229476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1959	100.0	380	10	US-09-936-883C-2
2	1959	100.0	380	12	US-10-126-052A-263
3	1959	100.0	380	14	US-60-443-566-3006
4	1959	100.0	380	14	US-60-453-444-4692
5	1959	100.0	380	14	US-60-465-241-4692
6	1477	75.4	380	10	US-09-936-883C-19
7	1458	74.4	368	10	US-09-936-883C-21
8	822	42.0	372	14	US-60-443-566-3717
9	822	42.0	372	14	US-60-455-444-5630
10	822	42.0	372	14	US-60-465-241-5630
11	822	42.0	392	12	US-10-198-070-62
12	820	41.9	392	12	US-10-198-070-67
13	819	41.8	392	12	US-10-198-070-70
14	796.5	40.7	379	12	US-10-294-025-946
15	769	39.3	394	12	US-10-258-951-59
16	763.5	39.0	405	2	PCT-US02-38526-416
17	763.5	39.0	405	14	US-60-443-566-4407
18	763.5	39.0	405	14	US-60-455-444-7462
19	763.5	39.0	405	14	US-60-465-241-7462

20	754.5	38.5	425	2	PCT-US02-072158-2	Sequence 2, Appl1
21	754.5	38.5	425	12	US-10-419-277-2	Sequence 2, Appl1
22	754.5	38.5	425	14	US-60-443-566-4406	Sequence 4406, Ap
23	754.5	38.5	425	14	US-60-455-444-7461	Sequence 7461, Ap
24	754.5	38.5	425	14	US-60-465-241-7461	Sequence 7461, Ap
25	753.5	38.5	425	2	PCT-US02-38526-408	Sequence 408, App
26	752	38.4	390	2	PCT-US02-072158-12	Sequence 12, Appl
27	752	38.4	390	2	PCT-US03-17409-175	Sequence 175, Appl
28	752	38.4	390	14	US-60-448-285-5	Sequence 5, Appl1
29	751	38.3	410	10	US-09-724-676A-57391	Sequence 57391, A
30	751	38.3	410	10	US-09-724-676A-57391	Sequence 57391, A
31	748	38.2	390	2	PCT-US02-38526-417	Sequence 417, App
32	748	38.2	390	2	PCT-US02-072158-13	Sequence 13, Appl
33	748	38.2	390	12	US-10-126-052A-465	Sequence 465, App
34	748	38.2	390	12	US-10-295-027-177	Sequence 177, App
35	748	38.2	390	12	US-10-295-027-794	Sequence 794, App
36	748	38.2	390	14	US-60-448-285-6	Sequence 6, Appl1
37	738.5	37.7	617	2	PCT-US02-072158-14	Sequence 14, Appl
38	738.5	37.7	617	12	US-10-294-025-947	Sequence 947, App
39	722.5	36.9	379	2	PCT-US03-12731-14	Sequence 14, Appl
40	722.5	36.9	379	2	PCT-US02-072158-9	Sequence 9, Appl1
41	722.5	36.9	379	14	US-60-448-285-3	Sequence 3, Appl1
42	722.5	36.9	379	14	US-60-453-135-8341	Sequence 8341, Ap
43	722.5	36.9	379	14	US-60-453-050-8341	Sequence 8341, Ap
44	722.5	36.9	379	14	US-60-455-444-4629	Sequence 4629, Ap
45	722.5	36.9	379	14	US-60-465-241-4629	Sequence 4629, Ap

ALIGNMENTS

RESULT 1
US-09-936-883C-2
Sequence 2, Application US/09936883C
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
TITLE OR INVENTION: Thereof
FILE REFERENCE: P2-101DIPCT
CURRENT APPLICATION NUMBER: US/09/936, 883C
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-936-883C-2

Query Match	100.0%	Score 1959;	DB 10;	Length 380;
Best Local Similarity	100.0%	Pred. No. 1,4e-181;		
Matches	380;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	MASLAANAEEFCNLEFREM	DGNGNGVNFESSLSLFAALALVRLGADDSISQIDKLHV	60
DB	1	MASLAANAEEFCNLEFREM	DGNGNGVNFESSLSLFAALALVRLGADDSISQIDKLHV	60
QY	61	NTASGYGSSNSGSGLOSLKRVFS	SDINASHKXVDLSVNGLFEXKYVGFHKDVECAEK	120
DB	61	NTASGYGSSNSGSGLOSLKRVFS	SDINASHKXVDLSVNGLFEXKYVGFHKDVECAEK	120
QY	121	LYDAKVRVDFPTNLEDTRRNIN	KVNEETHGKIKNYIGEGGSSAVMLVNAVYFKG	180
DB	121	LYDAKVRVDFPTNLEDTRRNIN	KVNEETHGKIKNYIGEGGSSAVMLVNAVYFKG	180
QY	181	WOSAFKTSSETINCHFKSPKCS	GKAVAMHDERKFNLSVIEDPSKILIELRYNGGINNYL	240
DB	181	WOSAFKTSSETINCHFKSPKCS	GKAVAMHDERKFNLSVIEDPSKILIELRYNGGINNYL	240

```

OY      241 LPENDLSEIENKLTFOINLMEWTNPRMTSKYVEVFFPOEKIEKNYEMKOYLALGKIDIF 300
        |||||||
Db      241 LPENDLSEIENKLTFOINLMEWTNPRMTSKYVEVFFPOEKIEKNYEMKOYLALGKIDIF 300
OY      301 DESKADLSGIASGRLYISRMHKSXYIEVTEGTEATAATGSNIYEKOLPOSTLFRADHP 360
        |||||||
Db      301 DESKADLSGIASGRLYISRMHKSXYIEVTEGTEATAATGSNIYEKOLPOSTLFRADHP 360
OY      361 FLEVIKRDIIILFSGKVSCP 380
        |||||||
Db      361 FLEVIKRDIIILFSGKVSCP 380

```

RESULT 2

```

US-10-126-052A-263
; Sequence 263, Application US/10126052A
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nalasha
; APPLICANT: Murray, Richard
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of Lung Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
; FILE REFERENCE: 018501-001300S
; CURRENT APPLICATION NUMBER: US/10/126, 052A
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/284,770
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/290,492
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/339,245
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 263
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-052A-263

```

```

Query Match      100.0%; Score 1959; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.4e-181;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 MASLAAANAEFCFNLREMDNDNGNGNVFFSSLSLFAALALVRLGAODDSLQIDKILHV 60
        |||||||
Db      1 MASLAAANAEFCFNLREMDNDNGNGNVFFSSLSLFAALALVRLGAODDSLQIDKILHV 60
OY      61 NTASGYGNSNSQSGIOSQKRVFSDINASHKDYDIYNGLFAEKVYGFHKDYICAEK 120
        |||||||
Db      61 NTASGYGNSNSQSGIOSQKRVFSDINASHKDYDIYNGLFAEKVYGFHKDYICAEK 120
OY      121 LYDAKVERVDFTNHLEDTRRNINKWVENETHGKIKRVIGEGGISSAVVVLNAVYFKGK 180
        |||||||
Db      121 LYDAKVERVDFTNHLEDTRRNINKWVENETHGKIKRVIGEGGISSAVVVLNAVYFKGK 180
OY      181 MOSAFKSETINCHFRSPKCGSKAVAMMOERKFNLSYEDPSMKLLEIRYNGIMYVL 240
        |||||||
Db      181 MOSAFKSETINCHFRSPKCGSKAVAMMOERKFNLSYEDPSMKLLEIRYNGIMYVL 240
OY      241 LPENDLSEIENKLTFOINLMEWTNPRMTSKYVEVFFPOEKIEKNYEMKOYLALGKIDIF 300
        |||||||
Db      241 LPENDLSEIENKLTFOINLMEWTNPRMTSKYVEVFFPOEKIEKNYEMKOYLALGKIDIF 300
OY      301 DESKADLSGIASGRLYISRMHKSXYIEVTEGTEATAATGSNIYEKOLPOSTLFRADHP 360
        |||||||
Db      301 DESKADLSGIASGRLYISRMHKSXYIEVTEGTEATAATGSNIYEKOLPOSTLFRADHP 360

```

```

OY      361 FLEVIKRDIIILFSGKVSCP 380
        |||||||
Db      361 FLEVIKRDIIILFSGKVSCP 380

```

RESULT 3

```

US-60-443-566-3006
; Sequence 3006, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443,566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 3006
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-3006

```

```

Query Match      100.0%; Score 1959; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.4e-181;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 MASLAAANAEFCFNLREMDNDNGNGNVFFSSLSLFAALALVRLGAODDSLQIDKILHV 60
        |||||||
Db      1 MASLAAANAEFCFNLREMDNDNGNGNVFFSSLSLFAALALVRLGAODDSLQIDKILHV 60
OY      61 NTASGYGNSNSQSGIOSQKRVFSDINASHKDYDIYNGLFAEKVYGFHKDYICAEK 120
        |||||||
Db      61 NTASGYGNSNSQSGIOSQKRVFSDINASHKDYDIYNGLFAEKVYGFHKDYICAEK 120
OY      121 LYDAKVERVDFTNHLEDTRRNINKWVENETHGKIKRVIGEGGISSAVVVLNAVYFKGK 180
        |||||||
Db      121 LYDAKVERVDFTNHLEDTRRNINKWVENETHGKIKRVIGEGGISSAVVVLNAVYFKGK 180
OY      181 MOSAFKSETINCHFRSPKCGSKAVAMMOERKFNLSYEDPSMKLLEIRYNGIMYVL 240
        |||||||
Db      181 MOSAFKSETINCHFRSPKCGSKAVAMMOERKFNLSYEDPSMKLLEIRYNGIMYVL 240
OY      241 LPENDLSEIENKLTFOINLMEWTNPRMTSKYVEVFFPOEKIEKNYEMKOYLALGKIDIF 300
        |||||||
Db      241 LPENDLSEIENKLTFOINLMEWTNPRMTSKYVEVFFPOEKIEKNYEMKOYLALGKIDIF 300
OY      301 DESKADLSGIASGRLYISRMHKSXYIEVTEGTEATAATGSNIYEKOLPOSTLFRADHP 360
        |||||||
Db      301 DESKADLSGIASGRLYISRMHKSXYIEVTEGTEATAATGSNIYEKOLPOSTLFRADHP 360
OY      361 FLEVIKRDIIILFSGKVSCP 380
        |||||||
Db      361 FLEVIKRDIIILFSGKVSCP 380

```

RESULT 4

```

US-60-455-444-4692
; Sequence 4692, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 4692
; LENGTH: 380
; TYPE: PRT

```

ORGANISM: Homo sapiens
US-60-465-444-4692

Query Match 100.0%; Score 1959; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.4e-181;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
DB 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
OY 61 NTASGCGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLEAEKYYGFHKDYIECAEK 120
DB 61 NTASGCGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLEAEKYYGFHKDYIECAEK 120
OY 121 LYDAKVERVDFTNHLLEDTNRINKWVENETHGKIKVYIGGGJSSSAVMVLVNAVYFKG 180
DB 121 LYDAKVERVDFTNHLLEDTNRINKWVENETHGKIKVYIGGGJSSSAVMVLVNAVYFKG 180
OY 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240
OY 241 LPENDLSEIENKLTFOULMEWTNPRMTSKYVEVFPOFKIEKNYEKKOYLRLGLKIDIF 300
DB 241 LPENDLSEIENKLTFOULMEWTNPRMTSKYVEVFPOFKIEKNYEKKOYLRLGLKIDIF 300
OY 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
OY 361 FLFVIRKDDIILFSGKVCSP 380
DB 361 FLFVIRKDDIILFSGKVCSP 380

RESULT 5

US-60-465-241-4692
Sequence 4692, Application US/60465241
GENERAL INFORMATION:
APPLICANT: CARGILH, Michele
APPLICANT: BEGOVICH, Ann
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001468
CURRENT APPLICATION NUMBER: US/60/465,241
NUMBER OF SEQ ID NOS: 258418
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4692
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-60-465-241-4692

Query Match 100.0%; Score 1959; DB 14; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.4e-181;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
DB 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
OY 61 NTASGCGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLEAEKYYGFHKDYIECAEK 120
DB 61 NTASGCGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLEAEKYYGFHKDYIECAEK 120
OY 121 LYDAKVERVDFTNHLLEDTNRINKWVENETHGKIKVYIGGGJSSSAVMVLVNAVYFKG 180
DB 121 LYDAKVERVDFTNHLLEDTNRINKWVENETHGKIKVYIGGGJSSSAVMVLVNAVYFKG 180
OY 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240

DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240
OY 241 LPENDLSEIENKLTFOULMEWTNPRMTSKYVEVFPOFKIEKNYEKKOYLRLGLKIDIF 300
DB 241 LPENDLSEIENKLTFOULMEWTNPRMTSKYVEVFPOFKIEKNYEKKOYLRLGLKIDIF 300
OY 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
OY 361 FLFVIRKDDIILFSGKVCSP 380
DB 361 FLFVIRKDDIILFSGKVCSP 380

RESULT 6

US-09-936-883C-19
Sequence 19, Application US/09936883C
GENERAL INFORMATION:
APPLICANT: MIYATA, Toshio
TITLE OF INVENTION: A Method for Detecting Megsin Protein and Use
TITLE OF INVENTION: thereof
FILE REFERENCE: F2-101DIPCT
CURRENT APPLICATION NUMBER: US/09/936, 883C
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: JP 1999-75305
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: JP 1999-306623
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 380
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc_feature
LOCATION: 51, 94
OTHER INFORMATION: Xaa is unknown.
US-09-936-883C-19

Query Match 75.4%; Score 1477; DB 10; Length 380;
Best Local Similarity 73.4%; Pred. No. 9e-135;
Matches 279; Conservative 56; Mismatches 45; Indels 0; Gaps 0;

OY 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
DB 1 MASLAANAEEFCNLFREMDNNGNNGVFFSSLSLFAALALVRLGADDSLSQIDKLHV 60
OY 61 NTASGCGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLEAEKYYGFHKDYIECAEK 120
DB 61 NTASGCGNSSNSGLOSOLKRVFSDINASHKDYDLSIVNGLEAEKYYGFHKDYIECAEK 120
OY 121 LYDAKVERVDFTNHLLEDTNRINKWVENETHGKIKVYIGGGJSSSAVMVLVNAVYFKG 180
DB 121 LYDAKVERVDFTNHLLEDTNRINKWVENETHGKIKVYIGGGJSSSAVMVLVNAVYFKG 180
OY 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240
DB 181 WSAFTKSETINCHFKSPKCSGKAVAMHDERKFNLSVIDPMSKILELRNGGINNYVL 240
OY 241 LPENDLSEIENKLTFOULMEWTNPRMTSKYVEVFPOFKIEKNYEKKOYLRLGLKIDIF 300
DB 241 LPENDLSEIENKLTFOULMEWTNPRMTSKYVEVFPOFKIEKNYEKKOYLRLGLKIDIF 300
OY 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
DB 301 DESKADLSGIASGRLYISRMHKSYLEVEEGTEATAAGSNIVEKQLPOSTLFRADHP 360
OY 361 FLFVIRKDDIILFSGKVCSP 380
DB 361 FLFVIRKDDIILFSGKVCSP 380

LENGTH: 392
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-198-070-67

Query Match 41.9%; Score 820; DB 12; Length 392;

Best Local Similarity 43.1%; Pred. No. 6.2e-71;

Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

QY 1 MASLAANAFECNLFREDDNGNGNVFESSLSLEPAALAYRLGAODDSLQIDKILHY 60
 DB 1 MGSLSLANVEFLDVKELNSNIGDNIFSSLSLYALSMVLLGARGETAEQLEKVLHF 60
 QY 61 -----NTASGYGNSN-SOSG-LQSOLKRVFSDINASHKDYDLSYNGLEFAEVYGFHKD 113
 DB 61 SHTVDSLKPGFKDSPCSOAGRIHSEGVFESQINOPDSNCTLSIANRLYGTMAFHQ 120
 QY 114 YIECAKLDKAKVERDFTNHLNEDTRNINKWYENETHGKIKVIEGGISSAVAVLVN 173
 DB 121 YLSCSEKWOARLQYVDFEOSTETETRKMINAWENTNGKVANLFGKSTIDPSSVAVLVN 180
 QY 174 AVFEKQWOSAFETSETINCHFKSPKSCGKAVAMHOERKFNLSVIEDPSMKITELRY-N 232
 DB 181 TIFKQORQNKFPVRETVSPFOLSEGNVTEMMQIGTFKLAFAVEKPOQVLELRYN 240
 QY 233 GGIMYVLLPEN--DLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPOFKIEKNYEMK 290
 DB 241 NKLSMITLLPVGIANLKQIEKQNSGFHEWTSNNMEREVEVHLPRFKLEIKYELNSL 300
 QY 291 LRALGLKDFDESKADLSGASGRLYISRMHKSITYEVEETETATGSIYVOKLP 350
 DB 301 LKPLGVTDLFNOVKADLSGMSPTKGLYLKAIHKSIVLDVSEGETAAATGDSIAVKSIP 360
 QY 351 OSTLFRAHDFLEFVIR--KDDIILFSGKVS 380
 DB 361 MRAQFKANHPLEFIRHTHTNTILFCGKLASP 392

RESULT 13

US-10-198-070-70

Sequence 70, Application US/10198070

GENERAL INFORMATION:

APPLICANT: AVERBACK, PAUL

APPLICANT: GEMMELL, JACK

TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER

TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF

FILE REFERENCE: 59003.000008

CURRENT APPLICATION NUMBER: US/10/198, 070

CURRENT FILING DATE: 2002-07-19

PRIOR APPLICATION NUMBER: 60/306, 161

PRIOR FILING DATE: 2001-07-19

PRIOR FILING DATE: 2001-07-19

PRIOR FILING DATE: 2001-07-19

PRIOR APPLICATION NUMBER: 60/331, 477

PRIOR FILING DATE: 2001-11-16

NUMBER OF SEQ ID NOS: 125

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 70

LENGTH: 392

TYPE: PRT

ORGANISM: Homo sapiens

US-10-198-070-70

Query Match 41.8%; Score 819; DB 12; Length 392;

Best Local Similarity 43.1%; Pred. No. 7.7e-71;

Matches 169; Conservative 80; Mismatches 131; Indels 12; Gaps 6;

QY 1 MASLAANAFECNLFREDDNGNGNVFESSLSLEPAALAYRLGAODDSLQIDKILHY 60
 DB 1 MGSLSLANVEFLDVKELNSNIGDNIFSSLSLYALSMVLLGARGETAEQLEKVLHF 60
 QY 61 -----NTASGYGNSN-SOSG-LQSOLKRVFSDINASHKDYDLSYNGLEFAEVYGFHKD 113

DB 61 SHTVDSLKPGFKDSPCSOAGRIHSEGVFESQINOPDSNCTLSIANRLYGTMAFHQ 120
 QY 114 YIECAKLDKAKVERDFTNHLNEDTRNINKWYENETHGKIKVIEGGISSAVAVLVN 173
 DB 121 YLSCSEKWOARLQYVDFEOSTETETRKMINAWENTNGKVANLFGKSTIDPSSVAVLVN 180
 QY 174 AVFEKQWOSAFETSETINCHFKSPKSCGKAVAMHOERKFNLSVIEDPSMKITELRY-N 232
 DB 181 TIFKQORQNKFPVRETVSPFOLSEGNVTEMMQIGTFKLAFAVEKPOQVLELRYN 240
 QY 233 GGIMYVLLPEN--DLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPOFKIEKNYEMK 290
 DB 241 NKLSMITLLPVGIANLKQIEKQNSGFHEWTSNNMEREVEVHLPRFKLEIKYELNSL 300
 QY 291 LRALGLKDFDESKADLSGASGRLYISRMHKSITYEVEETETATGSIYVOKLP 350
 DB 301 LKPLGVTDLFNOVKADLSGMSPTKGLYLKAIHKSIVLDVSEGETAAATGDSIAVKSIP 360
 QY 351 OSTLFRAHDFLEFVIR--KDDIILFSGKVS 380
 DB 361 MRAQFKANHPLEFIRHTHTNTILFCGKLASP 392

RESULT 14

US-10-294-025-946

Sequence 946, Application US/10294025

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Scolik, John A.

APPLICANT: Kalos, Michael D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C29

CURRENT APPLICATION NUMBER: US/10/294, 025

CURRENT FILING DATE: 2002-11-12

NUMBER OF SEQ ID NOS: 1038

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 946

LENGTH: 379

TYPE: PRT

ORGANISM: Homo sapiens

US-10-294-025-946

Query Match 40.7%; Score 796.5; DB 12; Length 379;

Best Local Similarity 42.6%; Pred. No. 1.1e-68;

Matches 168; Conservative 77; Mismatches 120; Indels 29; Gaps 8;

QY 1 MASLAANAFECNLFREDDNGNGNVFESSLSLEPAALAYRLGAODDSLQIDKILHY 60
 DB 1 MGSLSLANVEFLDVKELNSNIGDNIFSSLSLYALSMVLLGARGETAEQLEKVLHF 60
 QY 61 -----NTASGYGNSN-SOSG-LQSOLKRVFSDINASHKDYDLSYNGLEFAEVYGFHKD 113
 DB 61 SHTVDSLKPGFKDSPCSOAGRIHSEGVFESQINOPDSNCTLSIANRLYGTMAFHQ 120
 QY 114 YIECAKLDKAKVERDFTNHLNEDTRNINKWYENETHGKIKVIEGGISSAVAVLVN 173
 DB 121 YLSCSEKWOARLQYVDFEOSTETETRKMINAWENTNGKVANLFGKSTIDPSSVAVLVN 180
 QY 174 AVFEKQWOSAFETSETINCHFKSPKSCGKAVAMHOERKFNLSVIEDPSMKITELRY 231
 DB 181 TIFKQORQNKFPVRETVSPFOLSEGNVTEMMQIGTFKLAFAVEKPOQVLELRY 240
 QY 233 GGIMYVLLPEN--DLSEIENKLTFFONLMEWTPRRMTSKYVEVFFPOFKIEKNYEMK 288
 DB 241 NKLSMITLLPVGIANLKQIEKQNSGFHEWTSNNMEREVEVHLPRFKLEIKYELNSL 300
 QY 291 LRALGLKDFDESKADLSGASGRLYISRMHKSITYEVEETETATGSIYVOKLP 348
 DB 301 LKPLGVTDLFNOVKADLSGMSPTKGLYLKAIHKSIVLDVSEGETAAATGDSIAVKSIP 360
 QY 349 LPOSTLFRAHDFLEFVIR--KDDIILFSGKVS 380

THIS PAGE BLANK (USPTO)


```

XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
PS Claim 7; Page 30-32; 47pp; English.
XX
CC Human MDF can be isolated from a culture of human epidermoid
CC carcinoma A431 cells in protein-free medium. The MDF stimulates
CC differentiation of megakaryocytes from myeloid cells in the presence
CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC for treatment of diseases involving a decrease in platelet number
CC (esp. thrombocytopaenia) such as occurs in bone marrow
CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kd by
CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 52; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 2
AA08254
ID AA08254 standard; Protein: 380 AA.
XX
AC AA08254;
XX
DT 14-JUL-1999 (first entry)
XX
DE Human megsin protein.
XX
KW Megsin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
KW human; rat; murine.
XX
OS Homo sapiens.
XX
PN MO9915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI; 1999-276983/23.
DR N-PSDB; AAX56712.
XX
PT Megsin protein expressed specifically in mesangial cells
XX
PS Claim 1; Page 62-64; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megsin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as Iga nephropathy.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 52; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 3
AAB24142
ID AAB24142 standard; Protein: 380 AA.
XX
AC AAB24142;
XX
DT 30-JAN-2001 (first entry)
XX
DE Human megsin protein sequence SEQ ID NO:2.
XX
KW Megsin; mesangium-predominant gene; seipin regulated; nephropathy;
KW Iga; Immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
OS Homo sapiens.
XX
PN WO200057189-A1.
XX
PD 28-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-JP01646.
XX
PR 19-MAR-1999; 99JP-0075305.
PR 28-OCT-1999; 99JP-0306623.
XX
PA (KURO/) KUROKAWA K.
PA (FUSO ) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI; 2000-611642/58.
DR N-PSDB; AAA99294.
XX
PT Evaluating renal function comprises assaying megsin protein in
PT biological sample
XX
PS Example 2; Page 66-69; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megsin protein in biological sample. Also
CC described are: (1) use of a anti-megsin protein antibody for diagnosing
CC renal function; and (2) a kit for detecting megsin protein comprising:
CC (a) anti-megsin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying megsin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence represents the human megsin
CC protein, which is given in the exemplification of the present invention.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 52; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHFK 9
Db 188 SETINCHFK 196

RESULT 4
AAB83075
ID AAB83075 standard; Protein: 380 AA.
XX
AC AAB83075;

```

XX 10-JUL-2001 (first entry)
DT Human megsln protein.
XX
DE
XX
KW Human; megsln; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening.
XX
OS Homo sapiens.
XX
PN WO200124628-A1.
XX
PD 12-APR-2001.
XX
PE 06-OCT-2000; 2000WO-JP06988.
XX
PR 06-OCT-1999; 99JP-0285736.
XX
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
XX WPI: 2001-300136/31.
DR N-PSDB: AAF82438.
XX
XX
PT Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments -
XX
PS Example 4; Page 44-46; 62pp; Japanese.
XX
CC The present sequence is human megsln. The human megsln coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.
XX
SO Sequence 380 AA;
XX
QY Query Match 100.0%; Score 52; DB 22; Length 380;
Best Local Similarity 100.0%; Pred. NO. 0.092;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 SETINCHFK 9
DB 188 SETINCHFK 196
XX
RESULT 5
AAV08255
ID AAV08255 standard; Protein: 380 AA.
XX
XX
AC AAY08255;
XX
DT 14-JUL-1999 (first entry)
XX
XX Rat megsln protein.
DE
KW Megsln; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
KW human; rat; murine.
XX
XX Rattus rattus.
OS
XX
XX MO9915652-A1.
XX
XX 01-APR-1999.
XX
XX 22-SEP-1998; 98WO-JP04269.
XX
XX

PR 22-SEP-1997; 97JP-0275302.
XX
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
XX
PI Miyata T;
XX
XX WPI: 1999-276983/23.
DR N-PSDB: AAX56712.
XX
XX
PT Megsln protein expressed specifically in mesangial cells
PT
PS Claim 1; Page 69-72; 100pp; Japanese.
XX
XX
CC This invention describes the isolation of novel megsln nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as Iga nephropathy.
XX
XX
SO Sequence 380 AA;
XX
QY Query Match 75.0%; Score 39; DB 20; Length 380;
Best Local Similarity 55.6%; Pred. NO. 26;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
1 SETINCHFK 9
DB 188 SDTLSCFKR 196
XX
XX
RESULT 6
AAB24150
ID AAB24150 standard; Protein: 380 AA.
XX
XX
AC AAB24150;
XX
XX 30-JAN-2001 (first entry)
DT
DE Rat megsln protein sequence SEQ ID NO:19.
XX
XX
KW Megsln; mesangium-predominant gene; serpin regulated; nephropathy;
KW Iga; immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
XX
OS Rattus norvegicus.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 51
FT Misc-difference /note- "unspecified"
FT Misc-difference 94 /note- "unspecified"
XX
XX
PN WO200057189-A1.
XX
XX
PD 28-SEP-2000.
XX
XX
PE 17-MAR-2000; 2000WO-JP01646.
XX
XX
PR 19-MAR-1999; 99JP-0075305.
PR 28-OCT-1999; 99JP-0306623.
XX
XX
PA (KURO/) KUROKAWA K.
PA (FUSO) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
XX
PI Miyata T;
XX
XX
DR WPI: 2000-611642/58.
DR N-PSDB: AAC55238.
XX
XX
PT Evaluating renal function comprises assaying megsln protein in
PT biological sample -

xx	Example 2: Page 81-84; 93pp; Japanese.
ps	
xx	
xx	The present invention describes a method for evaluating renal function.
cc	The method comprises assaying mesgin protein in biological sample. Also
cc	described are: (1) use of an antimargin protein antibody for diagnosing
cc	renal function; and (2) a kit for detecting mesgin protein comprising:
cc	(a) antimargin protein antibody attached to solid magnetic particles;
cc	(b) direct or indirect fixing for the antibody to the particles; and
cc	(c) a magnet. The process is useful for evaluating renal function and
cc	diagnosing renal disorders by assaying mesgin protein in biological
cc	samples (preferably urine or blood). The process is reproducible and
cc	gives accurate results. The present sequence represents the rat mesgin
cc	protein, which is given in the exemplification of the present invention.
xx	
xx	Sequence 380 AA:
QY	1 SETINCHK 9
DB	188 SDTLSCHFR 196
RESULT 7	
AA664286	standard; Protein: 380 AA.
AC	AA664286;
XX	
XX	21-SEP-2001 (first entry)
DE	Rat mesgin protein.
XX	
KW	Rat; mesgin; renal mesangial cell; mesangium proliferative nephritis.
XX	
OS	Rattus norvegicus.
XX	
PN	WO200148019-A1.
XX	
PD	05-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-JP09251.
PR	28-DEC-1999; 99JP-0373677.
XX	
PA	(KUROO/) KUROKAWA K.
PA	(MIYA/) MIYATA T.
PI	Miyata T;
XX	
DR	WPI: 2001-425651/45.
DR	N-PSDB; AAH48181.
XX	
PT	New antibody recognizing a partial sequence of rat mesgin protein for
PT	diagnosis of mesangium proliferative nephritis -
XX	
PS	Disclosure: Page 54-56; 63pp; Japanese.
XX	
CC	The present invention relates to a novel antibody which recognises a
CC	peptide consisting of residues 341-354 of rat mesgin protein. The present
CC	sequence is the protein sequence for rat mesgin, which was used in the
CC	present invention. Mesgin is highly expressed in renal mesangial cells
CC	and its level is elevated in mesangium proliferative nephritis. Assay of
CC	the serum or urine level using the antibody is therefore indicative of
CC	this type of disorder.
XX	
SQ	Sequence 380 AA:
Query Match	75.0%; Score 39; DB 22; Length 380;
Best Local Similarity	55.6%; Pred. No. 26;
Matches	5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY	1 SETINCHK 9
DB	188 SDTLSCHFR 196
RESULT 8	
AA83076	standard; Protein: 380 AA.
AC	AA83076;
XX	
XX	10-JUL-2001 (first entry)
DE	Rat mesgin protein.
XX	
KW	Rat; mesgin; mesangial cell proliferative nephritis; nephrotropic;
KW	transgenic mouse; glomerular disease; animal model; drug screening.
XX	
OS	Rattus norvegicus.
XX	
PN	WO200124628-A1.
XX	
PD	12-APR-2001.
XX	
PF	06-OCT-2000; 2000WO-JP06988.
XX	
PR	06-OCT-1999; 99JP-0285736.
XX	
PA	(KUROO/) KUROKAWA K.
PA	(MIYA/) MIYATA T.
PI	Miyata T;
XX	
DR	WPI: 2001-300136/31.
DR	N-PSDB; AAF82439.
XX	
PT	Mouse model for mesangial cell proliferative nephritis for development
PT	and screening of new treatments -
XX	
PS	Disclosure: Page 48-50; 62pp; Japanese.
XX	
CC	The present sequence is rat mesgin. The human mesgin coding
CC	sequence may be introduced into a mouse to produce an animal model of
CC	mesangial cell proliferative nephritis. The symptoms include
CC	enlargement of the mesangial base region, sedimentation of an immune
CC	complex and an increase in mesangial cells. The animal model is useful
CC	for analysing the pathology of chronic glomerular diseases and for
CC	screening compositions for prevention and treatment of the diseases.
CC	Highly uniform models can be made easily and in large numbers using
CC	this method.
XX	
SQ	Sequence 380 AA:
Query Match	75.0%; Score 39; DB 22; Length 380;
Best Local Similarity	55.6%; Pred. No. 26;
Matches	5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY	1 SETINCHK 9
DB	188 SDTLSCHFR 196
RESULT 9	
ABB62242	standard; Protein: 199 AA.
ID	ABB62242
XX	
AC	ABB62242;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 13518.

XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PM WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, LI PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL06345.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 SQ Disclosure: SEQ ID NO 13518; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 199 AA:
 Query Match 73.1%; Score 38; DB 22; Length 199;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TINCHF 8
 Db 3 TINCHF 8
 RESULT 10
 ID ABB60820 standard; Protein; 208 AA.
 XX
 AC ABB60820;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 9252.
 XX
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PM WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, LI PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL04923.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 SQ Disclosure: SEQ ID NO 9252; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 208 AA:
 Query Match 73.1%; Score 38; DB 22; Length 208;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TINCHF 8
 Db 3 TINCHF 8
 RESULT 11
 ID AAR57107 standard; peptide; 9 AA.
 XX
 AC AAR57107;
 XX
 DT 16-AUG-1994 (first entry)
 XX
 DE Human megakaryocyte differentiation factor peptide 1.
 XX
 KW Human megakaryocyte differentiation factor; MDF; thrombopoietin;
 KW haematopoietic stimulating factor; thrombocytopoietin; platelet;
 KW bone marrow transplantation; cancer chemotherapy.
 XX
 OS Homo sapiens.
 XX
 PM EP583884-A.
 PD 23-FEB-1994.
 XX
 PF 19-JUL-1993; 93EP-0305654.
 XX
 PR 17-JUL-1992; 92JP-0212305.
 PR 04-MAR-1993; 93JP-0067339.
 XX
 PA (SUNR) SUNTORY LTD.
 PA (TSUJ/) TSUJIMOTO M.
 PI Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;

PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
 XX WPI: 1994-058782/08.
 XX
 PT New megakaryocyte differentiation factor - isolated from human
 PT epidermoid carcinoma cells, used to treat conditions involving a
 PT decrease in platelets
 XX
 PS Claim 1; Page 18; 47pp; English.
 XX
 CC Human MDF (see AAR48379) can be isolated from a culture of human
 CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
 CC stimulates differentiation of megakaryocytes from myeloid cells
 CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
 CC making it useful for treatment of diseases involving a decrease
 CC in platelet number (esp. thrombocytopenia) such as occurs in bone
 CC marrow transplantation and in chemotherapy. MDF has mol wt
 CC 55-57KD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
 CC contains an amino acid sequence comprising at least one of the
 CC sequences AAR57107-R57115.
 XX
 SQ Sequence 9 AA:
 XX
 Query Match 71.2%; Score 37; DB 15; Length 9;
 Best Local Similarity 87.3%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ETINCPRK 9
 |||||
 Db 2 ETINCPRK 9
 2
 RESULT 12
 AAR79450
 ID AAR79450 standard; Protein: 320 AA.
 XX
 AC AAR79450;
 XX
 DT 19-JAN-1996 (first entry)
 XX
 DE Rat A3 adenosine receptor.
 XX
 KW A3 adenosine receptor; screening; identification; agonist;
 KW antagonist; drug; probe; detection; genetic disorder; RFLP;
 KW restriction fragment length polymorphism.
 XX
 OS Rattus rattus.
 XX
 PN US5441883-A.
 XX
 PD 15-AUG-1995.
 XX
 PF 03-MAR-1992; 9205-0847563.
 XX
 PR 03-MAR-1992; 9205-0847563.
 PR 02-AUG-1993; 9305-0101435.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 PA
 XX Clavelli O, Zhou Q;
 PI
 XX WPI: 1995-292492/38.
 DR N-PSDB: AAO97312.
 XX
 PT DNA and protein sequence(s) for novel adenosine receptor, A3 - also
 PT recombinant vector contg the DNA, DNA probe(s), and cell culture
 PT transformed with the vector.
 XX
 PS Claim 1; Figure 1; 29pp; English.
 XX
 CC A3 adenosine receptors produced from cloned genes may be used to
 CC screen compounds for A3 adenosine receptor activity, or for
 CC determining the amount of adenosine-agonist or antagonist drug in a

CC solution. The DNA or its fragments may also be used as probes to
 CC determine tissue distribution of the receptors, to detect the
 CC presence or absence of the gene or in RFLP to detect genetic
 CC disorders.
 XX
 SQ Sequence 320 AA:
 XX
 Query Match 71.2%; Score 37; DB 16; Length 320;
 Best Local Similarity 55.6%; Pred. No. 52;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SETINCPRK 9
 |||||
 Db 163 SFTLSCPRK 171
 163
 RESULT 13
 ABP35638
 ID ABP35638 standard; Protein: 1030 AA.
 XX
 AC ABP35638;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Fungal ZBC protein sequence #64.
 XX
 KW Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein;
 KW antibacterial; beta-lactam; anti-hypercholesterolemic; lovastatin;
 KW mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
 KW angiotensin inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
 KW fungal toxin; cell surface receptor; plant growth regulator; pigment;
 KW insecticide; antineoplastic.
 XX
 OS Unidentified.
 XX
 PN WO200224865-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 19-SEP-2001; 2001WO-US29288.
 XX
 PR 19-SEP-2000; 2000US-233564P.
 XX
 XX (MICR-) MICROBIA INC.
 PA
 XX Holtzman D, Madden K, Maxon M, Sherman A;
 PI
 XX WPI: 2002-352005/38.
 DR N-PSDB: ABN79827.
 XX
 PT New method for improving the production of a secondary metabolite e.g.
 PT antineoplastic agent, ergot alkaloid from a fungus involves modulation
 PT of the expression of at least one zinc binuclear cluster protein gene
 XX
 XX Disclosure: SEQ ID 140; 49pp + sequence listing; English.
 XX
 CC The invention relates to improving the production of a secondary
 CC metabolite by a fungus. This involves modulating the expression of at
 CC least one ZBC (zinc binuclear cluster protein) gene in a manner to
 CC improve the yield of the secondary metabolite. Methods of the invention
 CC may be used for improving the production of the secondary metabolite e.g.
 CC antibacterial (such as beta-lactam), an anti-hypercholesterolemic (such
 CC as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),
 CC an ergot alkaloid (such as ergotamine), an angiotensin inhibitor (such
 CC as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,
 CC a fungal toxin, a modulator of cell surface receptor signalling, a plant
 CC growth regulator, a pigment, an insecticide, or an antineoplastic
 CC compound. The method results in a decrease in fermentor run-time, a
 CC decrease in the size of the fermentor required for the production of
 CC equivalent amounts of the secondary metabolite, or a decrease in the
 CC biomass required for the production, which translates into decreased
 CC waste that must be handled in downstream processing. The sequences given

CC In records ABP35575-ABP35722 represent ZBC proteins.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1030 AA;

Query Match 71.2%; Score 37; DB 23; Length 1030;
Best Local Similarity 55.6%; Pred No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHER 9

DB 689 SRVLCNHR 697

RESULT 14

AA009905
ID AA009905 standard; Protein; 64 AA.

XX
AC AA009905;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 23797.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

DR N-PSDB: AA189836.

XX Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PT disorders -

PS Claim 20: SEQ ID NO 23797; 1399pp + sequence listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 64 AA;

Query Match 69.2%; Score 36; DB 22; Length 64;

Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCH 7

DB 55 SETISCH 61

RESULT 15

ABP35700
ID ABP35700 standard; Protein; 863 AA.

XX
AC ABP35700;

DT 24-JUL-2002 (first entry)

XX Fungal ZBC protein sequence #126.

XX Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein;
XX antibacterial; beta-lactam; anti-hypercholesterolemic; lovastatin;
XX mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
XX angiogenesis inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
XX fungal toxin; cell surface receptor; plant growth regulator; pigment;
XX insecticide; antineoplastic.

OS Unidentified.

PN WO200224865-A2.

PD 28-MAR-2002.

PF 19-SEP-2001; 2001WO-US29288.

XX 19-SEP-2000; 2000US-233564P.

PA (MICR-) MICROBIA INC.

PI Holtzman D, Madden K, Maxon M, Sherman A;

XX WPI: 2002-352005/38.

DR N-PSDB: ABN79889.

XX New method for improving the production of a secondary metabolite e.g.
PT antineoplastic agent, ergot alkaloid from a fungus involves modulation
PT of the expression of at least one zinc binuclear cluster protein gene

PS Disclosure: SEQ ID 264; 49pp + sequence listing; English.

XX The invention relates to improving the production of a secondary
CC metabolite by a fungus. This involves modulating the expression of at
CC least one ZBC (zinc binuclear cluster protein) gene in a manner to
CC improve the yield of the secondary metabolite. Methods of the invention
CC may be used for improving the production of the secondary metabolite e.g.
CC as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),
CC an ergot alkaloid (such as ergotamine), an angiogenesis inhibitor (such
CC as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,
CC a fungal toxin, a modulator of cell surface receptor signaling, a plant
CC growth regulator, a pigment, an insecticide, or an antineoplastic
CC compound. The method results in a decrease in fermentor run-time, a
CC decrease in the size of the fermentor required for the production of
CC equivalent amounts of the secondary metabolite, or a decrease in the
CC biomass required for the production, which translates into decreased
CC waste that must be handled in downstream processing. The sequences given
CC in records ABP35575-ABP35722 represent ZBC proteins.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 863 AA;

Query Match 69.2%; Score 36; DB 23; Length 863;

Best Local Similarity 62.5%; Pred. No. 2.3e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ETINCHFK 9
 : 1: 1111
 Db 470 KTLOCHFK 477

Search completed: July 11, 2003, 11:54:17
 Job time : 6.16424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 1.60915 seconds

(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-1

Perfect score: 52

Sequence: 1 SETINCHK 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6CTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6ackfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	9	2	US-08-611-977-1
2	52	100.0	380	2	US-08-472-659-34
3	52	100.0	380	2	US-08-474-661-34
4	52	100.0	380	2	US-08-611-977-34
5	37	71.2	9	2	US-08-472-659-1
6	37	71.2	9	2	US-08-474-661-1
7	37	71.2	320	2	US-08-101-435-4
8	35	67.3	481	5	PCR-US91-02166-13
9	35	67.3	858	4	US-07-956-483-17
10	35	67.3	915	4	US-08-472-240A-8
11	34	65.4	185	2	US-08-463-911-3
12	34	65.4	195	3	US-08-955-937A-4
13	34	65.4	195	3	US-09-300-985-4
14	34	65.4	201	3	US-08-955-937A-2
15	34	65.4	201	3	US-09-300-985-2
16	34	65.4	236	4	US-09-140-804-6
17	34	65.4	804	4	US-09-067-091-2
18	34	65.4	1078	1	US-08-485-588-7
19	34	65.4	1078	1	US-08-484-751-7
20	34	65.4	1078	2	US-08-943-986-7
21	34	65.4	1078	3	US-08-943-986-7
22	34	65.4	1078	3	US-08-353-784-7
23	34	65.4	1078	3	US-08-484-719B-7
24	34	65.4	1078	4	US-08-484-159-7
25	34	65.4	1079	1	US-08-485-588-8
26	34	65.4	1079	1	US-08-484-565-8
27	34	65.4	1079	2	US-08-480-751-8

28	34	65.4	1079	2	US-08-943-986-8	Sequence 8, Appl1
29	34	65.4	1079	3	US-08-353-784-8	Sequence 8, Appl1
30	34	65.4	1079	3	US-08-484-719B-8	Sequence 8, Appl1
31	34	65.4	1079	4	US-08-484-159-8	Sequence 8, Appl1
32	34	65.4	1085	1	US-08-485-588-5	Sequence 5, Appl1
33	34	65.4	1085	1	US-08-484-565-5	Sequence 5, Appl1
34	34	65.4	1085	2	US-08-480-751-5	Sequence 5, Appl1
35	34	65.4	1085	2	US-08-943-986-5	Sequence 5, Appl1
36	34	65.4	1085	3	US-08-353-784-5	Sequence 5, Appl1
37	34	65.4	1085	3	US-08-484-719B-5	Sequence 5, Appl1
38	34	65.4	1085	4	US-08-484-159-5	Sequence 5, Appl1
39	34	65.4	1088	1	US-08-485-588-6	Sequence 6, Appl1
40	34	65.4	1088	1	US-08-484-565-6	Sequence 6, Appl1
41	34	65.4	1088	2	US-08-480-751-6	Sequence 6, Appl1
42	34	65.4	1088	2	US-08-943-986-6	Sequence 6, Appl1
43	34	65.4	1088	3	US-08-353-784-6	Sequence 6, Appl1
44	34	65.4	1088	3	US-08-484-719B-6	Sequence 6, Appl1
45	34	65.4	1088	4	US-08-484-159-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-611-977-1
; Sequence 1, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAICHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-1

Query Match 100.0%; Score 52; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
1111111111
DB 1 SETINCHFK 9

RESULT 2
US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5831030uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5831030uhlro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAICHI, Kozo

APPLICANT: YAMAGUCHI, No. 5831030oml

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-248

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-472-659-34
Query Match 100.0%; Score 52; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
1111111111
DB 188 SETINCHFK 196

RESULT 3
US-08-474-661-34

Sequence 34, Application US/08474661
Patent No. 5874253

GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5874253uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5874253uhlro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMAICHI, Kozo

APPLICANT: YAMAGUCHI, No. 5874253oml

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: George Mason Bldg., Washington & Prince Sts.

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER

REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 001560-204

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-474-661-34
Query Match 100.0%; Score 52; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHFK 9
1111111111
DB 188 SETINCHFK 196

RESULT 4

US-08-611-977-34
; Sequence 34, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROUOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611.977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-611-977-34

Query Match 100.0%; Score 52; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
DB 188 SETINCHK 196

RESULT 5
US-08-472-659-1
; Sequence 1, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROUOKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5831030uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5831030oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.659
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-067339
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-248
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-472-659-1

Query Match 71.2%; Score 37; DB 2; Length 9;
Best Local Similarity 87.5%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCHK 9
DB 2 ETINCHK 9

RESULT 6
US-08-474-661-1
; Sequence 1, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROUOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhiro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5874253oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: George Mason Bldg., Washington & Prince Sts.
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,661
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 5-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REA, TERESA STANER
;; REGISTRATION NUMBER: 30,427
;; REFERENCE/DOCKET NUMBER: 001560-204
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-6620
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-474-661-1

Query Match 71.2%; Score 37; DB 2; Length 9;
Best Local Similarity 87.5%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCHFK 9
Db 2 ETINCHFK 9

RESULT 7
US-08-101-435-4
; Sequence 4, Application US/08101435
; Patent No. 5441883
; GENERAL INFORMATION:
; APPLICANT: Civeall, Olivier.
; APPLICANT: Zhou, Qun-Yong
; TITLE OF INVENTION: A No. 5441883el Adenosine Receptor and Uses
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/101,435
; FILING DATE:

;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/847,563
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McDonnell, John J
;; REGISTRATION NUMBER: 26,949
;; REFERENCE/DOCKET NUMBER: 91,708
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-715-1000
;; TELEFAX: 312-715-1234
;; TELEX: 910-221-5317
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 320 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-101-435-4

Query Match 71.2%; Score 37; DB 1; Length 320;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHFK 9
Db 163 SSTLSCHFR 171

RESULT 8
PCT-US91-02166-13
; Sequence 13, Application PC/TUS9102166
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: HIV Envelope Polypeptides
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: pacin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02166
; FILING DATE: 19910401
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S.S.N. 07/504,772
; FILING DATE: 03-APRIL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Carolyn R.
; REGISTRATION NUMBER: 32,324
; REFERENCE/DOCKET NUMBER: 639
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-2614
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-02166-13

Query Match 67.3%; Score 35; DB 5; Length 481;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCH 8
11111
Db 140 ETINCOF 146

RESULT 9
US-07-956-483-17
; Sequence 17, Application US/07956483
; Patent No. 6261799
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: 9p160 VARIANT
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/956,483
; FILING DATE: 31-DEC-1992
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO 92/19742
; FILING DATE: 12-NOV-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: FR 91 05392
; FILING DATE: 02-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 017753-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-956-483-17
Query Match 67.3%; Score 35; DB 4; Length 858;
Best Local Similarity 85.7%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCH 8
11111
Db 159 ETINCOF 165

RESULT 10
US-08-472-240A-8
; Sequence 8, Application US/08472240A
; Patent No. 6284248
; GENERAL INFORMATION:
; APPLICANT: KIENY, Marie-Paule
; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
; TITLE OF INVENTION: 9p160 VARIANT
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,240A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/956,483
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 017753-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..883
; US-08-472-240A-8

Query Match 67.3%; Score 35; DB 4; Length 915;
Best Local Similarity 85.7%; Pred. No. 16+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCH 8
11111
Db 182 ETINCOF 188

RESULT 11
US-08-463-911-3
; Sequence 3, Application US/08463911
; Patent No. 5869330
; GENERAL INFORMATION:
; APPLICANT: Scherer, Philipp E.
; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,911
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI95-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-463-911-3

Query Match 65.4%; Score 34; DB 2; Length 185;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TINCHK 9
Db 51 TVNCHK 57

RESULT 12
US-08-955-937A-4
Sequence 4, Application US/08955937A
Patent No. 6020161
GENERAL INFORMATION:
APPLICANT: WU, SHUTIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNEH, ALEMSSED
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,937A
FILING DATE: 17-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,152
FILING DATE: 19-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-937A-4

Query Match 65.4%; Score 34; DB 3; Length 195;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SETINCHK 9
Db 31 SLTVQCHK 39

RESULT 13
US-09-300-985-4
Sequence 4, Application US/09300985A
Patent No. 6232441
GENERAL INFORMATION:
APPLICANT: WU, SHUTIAN
APPLICANT: SWEET, RAYMOND W.
APPLICANT: TRUNEH, ALEMSSED
APPLICANT: HURLE, MARK ROBERT
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE
FILE REFERENCE: GH-70228-1
CURRENT APPLICATION NUMBER: US/09/300,985A
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: US 08/955,937
EARLIER FILING DATE: 1997-10-22
EARLIER APPLICATION NUMBER: US 60/056,152
EARLIER FILING DATE: 1997-08-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 195
TYPE: PRT
ORGANISM: HOMO SAPIENS
FEATURE:
NAME/KEY: UNSURE
LOCATION: (122)(170)(184)
US-09-300-985-4

Query Match 65.4%; Score 34; DB 4; Length 195;
Best Local Similarity 55.6%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SETINCHK 9
Db 31 SLTVQCHK 39

RESULT 14
US-08-955-937A-2
Sequence 2, Application US/08955937A
Patent No. 6020161
GENERAL INFORMATION:
APPLICANT: WU, SHUTIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNEH, ALEMSSED
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,937A
FILING DATE: 17-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,152

FILING DATE: 19-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-955-937A-2

Query Match 65.4%; Score 34; DB 3; Length 201;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
| | : | : |
DB 31 SLTVQCHYK 39

RESULT 15
US-09-300-985-2
Sequence 2, Application US/09300985A
Patent No. 6232441
GENERAL INFORMATION:
APPLICANT: WU, SHUTIAN
APPLICANT: SWEET, RAYMOND W.
APPLICANT: TRUNER, ALEKSEED
APPLICANT: HURLE, MARK ROBERT
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE
FILE REFERENCE: GH-70228-1
CURRENT APPLICATION NUMBER: US/09/300,985A
CURRENT FILING DATE: 1999-04-28
EARLIER APPLICATION NUMBER: US 08/955,937
EARLIER FILING DATE: 1997-10-22
EARLIER APPLICATION NUMBER: US 60/056,152
EARLIER FILING DATE: 1997-08-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 201
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-300-985-2

Query Match 65.4%; Score 34; DB 4; Length 201;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
| | : | : |
DB 31 SLTVQCHYK 39

Search completed: July 11, 2003, 12:02:04
Job time : 2.60915 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 2.63825 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-1
Perfect score: 52
Sequence: 1 SETINCHFK 9

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppa/US07_NEM_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PC10_NEM_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/PC10_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEM_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	9	US-10-091-442-1	Sequence 1, Appli
2	52	100.0	9	US-09-140-719-1	Sequence 1, Appli
3	52	100.0	380	US-10-091-442-34	Sequence 34, Appli
4	52	100.0	380	US-09-140-719-34	Sequence 34, Appli
5	35	67.3	90	US-09-764-881-113	Sequence 616, App
6	35	67.3	90	US-09-764-881-113	Sequence 616, App
7	34	65.4	201	US-10-174-590-608	Sequence 608, App
8	34	65.4	201	US-10-174-590-608	Sequence 608, App
9	34	65.4	201	US-10-174-590-608	Sequence 608, App
10	34	65.4	201	US-10-174-590-608	Sequence 608, App
11	34	65.4	201	US-10-174-590-608	Sequence 608, App
12	34	65.4	201	US-10-174-590-608	Sequence 608, App
13	34	65.4	201	US-10-174-590-608	Sequence 608, App
14	34	65.4	201	US-10-174-590-608	Sequence 608, App
15	34	65.4	201	US-10-174-590-608	Sequence 608, App
16	34	65.4	201	US-10-174-590-608	Sequence 608, App
17	34	65.4	201	US-10-174-590-608	Sequence 608, App
18	34	65.4	201	US-10-174-590-608	Sequence 608, App
19	34	65.4	201	US-10-174-590-608	Sequence 608, App

20	34	65.4	201	US-10-174-579-608	Sequence 608, App
21	34	65.4	201	US-10-174-582-608	Sequence 608, App
22	34	65.4	201	US-10-174-588-608	Sequence 608, App
23	34	65.4	201	US-10-175-739-608	Sequence 608, App
24	34	65.4	201	US-10-175-740-608	Sequence 608, App
25	34	65.4	201	US-10-175-743-608	Sequence 608, App
26	34	65.4	201	US-10-176-488-608	Sequence 608, App
27	34	65.4	201	US-10-176-492-608	Sequence 608, App
28	34	65.4	201	US-10-176-492-608	Sequence 608, App
29	34	65.4	201	US-10-176-750-608	Sequence 608, App
30	34	65.4	201	US-10-176-985-608	Sequence 608, App
31	34	65.4	201	US-10-176-987-608	Sequence 608, App
32	34	65.4	201	US-10-176-991-608	Sequence 608, App
33	34	65.4	201	US-10-176-992-608	Sequence 608, App
34	34	65.4	201	US-10-176-993-608	Sequence 608, App
35	34	65.4	201	US-10-184-658-608	Sequence 608, App
36	34	65.4	201	US-10-173-635-608	Sequence 608, App
37	34	65.4	201	US-10-173-697-608	Sequence 608, App
38	34	65.4	201	US-10-173-705-608	Sequence 608, App
39	34	65.4	201	US-10-174-576-608	Sequence 608, App
40	34	65.4	201	US-10-174-585-608	Sequence 608, App
41	34	65.4	201	US-10-174-586-608	Sequence 608, App
42	34	65.4	201	US-10-175-747-608	Sequence 608, App
43	34	65.4	201	US-10-176-481-608	Sequence 608, App
44	34	65.4	201	US-10-176-485-608	Sequence 608, App
45	34	65.4	201	US-10-176-487-608	Sequence 608, App

ALIGNMENTS

RESULT 1
US-10-091-442-1
Sequence 1, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUTSUMOTO, Masafumi
IWASA, Fuyuki
TSUTSUMOTO, No. US20020164711A1
NAKAZATO, Hiroshi
MURA, Kenju
ISHIDA, No. US20020164711A1
KURIHARA, Tatsuya
YAMAGUCHI, Kozi
YAMAGUCHI, No. US20020164711A1
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF INVENTIONS: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-Aug-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-Jun-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-Jul-1993
APPLICATION NUMBER: JP 4-213305
FILING DATE: 17-Jul-1992
APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-091-442-1

Query Match 100.0%; Score 52; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHK 9
Db 1 SETINCHK 9

RESULT 2
US-09-140-719-1
Sequence 1, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMACHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-1

Query Match 100.0%; Score 52; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SETINCHK 9
Db 1 SETINCHK 9

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhlro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMACHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-MAR-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 380 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
 US-10-091-442-34

Query Match 100.0%; Score 52; DB 9; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.09;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 DB 188 SETINCHK 196

RESULT 4

; Sequence 34, Application US/09140719
 ; Patent No. US20010026931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TSUJIMOTO, Masafumi
 ; APPLICANT: IWASHI, Fuyuki
 ; APPLICANT: TSURUOKA, No. US20010026931A1
 ; APPLICANT: NAKAZATO, Hiroshi
 ; APPLICANT: MIURA, Kenju
 ; APPLICANT: ISHIDA, No. US20010026931A1
 ; APPLICANT: KURIHARA, Tatsuya
 ; APPLICANT: YAMAGUCHI, Kozo
 ; APPLICANT: YAMAGUCHI, No. US20010026931A1
 ; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/140,719
 ; FILING DATE: 08-AUG-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/474,661
 ; FILING DATE: 07-JUN-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/091,028
 ; FILING DATE: 14-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 4-212305
 ; FILING DATE: 17-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-067339
 ; FILING DATE: 04-MAR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McGowan, Malcolm K.
 ; REGISTRATION NUMBER: 39,300
 ; REFERENCE/DOCKET NUMBER: 001560-247
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 380 amino acids
 ; TYPE: amino acid

; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-140-719-34

Query Match 100.0%; Score 52; DB 10; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.09;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 DB 188 SETINCHK 196

RESULT 5

; Sequence 113, Application US/09764881
 ; Publication No. US20030125246A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT207
 ; CURRENT APPLICATION NUMBER: US/09/764,881
 ; CURRENT FILING DATE: 2001-01-17
 ; PRIOR application data removed - refer to PAM or file wrapper
 ; NUMBER OF SEQ ID NOS: 192
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 113
 ; LENGTH: 90
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (20)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-764-881-113

Query Match 67.3%; Score 35; DB 9; Length 90;
 Best Local Similarity 77.8%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 DB 69 SETINCHK 77

RESULT 6

; Sequence 616, Application US/09764853
 ; Patent No. US20020090672A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PJ206
 ; CURRENT APPLICATION NUMBER: US/09/764,853
 ; CURRENT FILING DATE: 2001-01-17
 ; PRIOR application data removed - consult PAM or file wrapper
 ; NUMBER OF SEQ ID NOS: 939
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 616
 ; LENGTH: 90
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (20)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-09-764-853-616

Query Match 67.3%; Score 35; DB 10; Length 90;
 Best Local Similarity 77.8%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9

Db 69 SETINCVLK 77

RESULT 7

US-10-174-590-608
; Sequence 608, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 31 SLTVQCHXK 39

RESULT 8

US-10-176-758-608
; Sequence 608, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 SETINCHK 9
Db 31 SLTVQCHXK 39

RESULT 9

US-10-175-737-608
; Sequence 608, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
Db 31 SLTVQCHXK 39

RESULT 10

US-10-173-706-608
; Sequence 608, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 1 1:11:1
DB 31 SLTVOCHK 39

RESULT 11
US-10-175-738-608
; Sequence 608, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 1 1:11:1
DB 31 SLTVOCHK 39

RESULT 12
US-10-175-752-608
; Sequence 608, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 1 1:11:1
DB 31 SLTVOCHK 39

RESULT 13
US-10-176-482-608
; Sequence 608, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
 1 1:11:1
DB 31 SLTVOCHK 39

RESULT 14
US-10-176-757-608
; Sequence 608, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
| : | : | : |
DB 31 SLTVQCHK 39

RESULT 15
US-10-176-913-608
; Sequence 608, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 608
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-608

Query Match 65.4%; Score 34; DB 9; Length 201;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCHK 9
| : | : | : |
DB 31 SLTVQCHK 39

Search completed: July 11, 2003, 12:37:31
Job time : 3.63825 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 1.81497 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-1

Perfect score: 52

Sequence: 1 SETINCHFR 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: PIR73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	37	71.2	320	2	S17177
2	37	71.2	320	2	A46152
3	37	71.2	337	2	S68678
4	37	71.2	1030	2	S57380
5	36	69.2	863	2	S38140
6	35	67.3	118	2	B64028
7	35	67.3	203	2	I48054
8	35	67.3	267	2	I49053
9	35	67.3	267	2	I55686
10	35	67.3	280	2	I49052
11	35	67.3	280	2	D84968
12	35	67.3	426	2	F95058
13	35	67.3	522	2	H97927
14	35	67.3	858	1	VCLJG2
15	35	67.3	1173	2	T25539
16	35	67.3	2095	2	S29529
17	34	65.4	82	2	T10149
18	34	65.4	133	2	I46632
19	34	65.4	136	2	I46635
20	34	65.4	138	2	I46634
21	34	65.4	145	2	I46633
22	34	65.4	215	2	C48150
23	34	65.4	231	1	HLCHBL
24	34	65.4	256	2	T28853
25	34	65.4	389	2	F81211
26	34	65.4	389	2	A81787
27	34	65.4	779	2	A86862
28	34	65.4	798	2	C95202
29	34	65.4	798	2	C98069

30	34	65.4	1078	2	A56715	calcium receptor (
31	34	65.4	1079	2	I59362	calcium/polyvalent
32	34	65.4	1085	2	S40476	Ca(2+)-sensing rec
33	34	65.4	1088	2	B56715	calcium receptor (
34	34	63.5	59	2	A62023	hypothetical prote
35	33	63.5	129	2	B48909	adenosine receptor
36	33	63.5	131	2	T29743	hypothetical prote
37	33	63.5	169	2	T33949	hypothetical prote
38	33	63.5	226	2	T39298	hypothetical prote
39	33	63.5	288	2	H75069	cytochrome-c3 hyd
40	33	63.5	298	2	T33220	hypothetical prote
41	33	63.5	313	2	A48903	beta-lactamase - P
42	33	63.5	320	2	E71696	rare lipoprotein A
43	33	63.5	348	2	D89102	protein F25E5.2 (l
44	33	63.5	347	2	T20472	hypothetical prote
45	33	63.5	367	2	T18839	hypothetical prote

ALIGNMENTS

RESULT 1
S17177
Probable G protein-coupled receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 04-Dec-1992 #sequence-revision 04-Dec-1992 #text-change 29-Sep-1999
C:Accession: S17177
R:Meyerhof, W.; Mueller-Brechlin, R.; Richter, D.
FEBS Lett. 284, 155-160, 1991
A:Title: Molecular cloning of a novel putative G-protein coupled receptor expressed d
A:Reference number: S17177; MUID:91285122; PMID:1647979
A:Accession: S17177
A:Molecule type: mRNA
A:Residues: 1-320 <MEY>
A:Cross-references: EMBL:X59249; NID:g56307; PIDN:CAA41937.1; PID:g56308
C:Superfamily: adenosine receptor A1
C:Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; thiolester bond; t
F:17-79/Domain: transmembrane #status predicted <TM1>
F:51-74/Domain: transmembrane #status predicted <TM2>
F:84-108/Domain: transmembrane #status predicted <TM3>
F:129-150/Domain: transmembrane #status predicted <TM4>
F:180-200/Domain: transmembrane #status predicted <TM5>
F:234-257/Domain: transmembrane #status predicted <TM6>
F:264-286/Domain: transmembrane #status predicted <TM7>
F:4,5,42,162,213/Binding site: carbonylate (asn) (covalent) #status predicted
F:302/Binding site: palmitate (cys) (covalent) #status predicted

Query Match
Best local similarity 55.6%; Pred. No. 25;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SETINCHFR 9
Db 163 SSTLSCHFR 171

RESULT 2
A46152
A3 adenosine receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Sep-1993 #sequence-revision 18-Nov-1994 #text-change 29-Sep-1999
C:Accession: A46152
R:Zhong, O.Y.; Li, C.; Olah, M.E.; Johnson, R.A.; Stiles, G.L.; Clivell, O.
Proc. Natl. Acad. Sci. U.S.A. 89, 7432-7436, 1992
A:Title: Molecular cloning and characterization of an adenosine receptor: the A3 aden
A:Reference number: A46152; MUID:92366475; PMID:1323836
A:Accession: A46152
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-320 <ZHO>
A:Cross-references: GB:M94152; NID:g202724; PIDN:AAA40680.1; PID:g460332
A:Experimental source: brain
A>Note: sequence inconsistent with the nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIN:110661, NCBIPI:110662)
C:Superfamily: adenosine receptor A1
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 71.2%; Score 37; DB 2; Length 320;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SETINCHFK 9
Db 163 STLSCHFR 171

RESULT 3

adenosine receptor A31 - rat
C:Species: Rattus sp. (rat)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68678
R:Satjadj, F.G.; Boyle, D.L.; Domingo, R.C.; Firestein, G.S.
FEBS Lett. 382, 125-129, 1996
A:Title: cDNA cloning and characterization of A31, an alternatively spliced rat A3 adenc
A:Reference number: S68678; MUID:96196578; PMID:8612733
A:Accession: S68678
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-337 <SAB>
A:Cross-references: EMBL:X93219; NID:91297000; PIDN:CA63702.1; PID:91297001
C:Superfamily: adenosine receptor A1
C:Keywords: alternative splicing

Query Match 71.2%; Score 37; DB 2; Length 337;
Best Local Similarity 55.6%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SETINCHFK 9
Db 180 STLSCHFR 188

RESULT 4

probable membrane protein YOL089c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein 00938
C:Species: Saccharomyces cerevisiae
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-Apr-2002
C:Accession: S57380; S66783; S50416
R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
Yeast 11, 975-986, 1995
A:Title: A 29,425 kb segment on the left arm of yeast chromosome XV contains more than t
A:Reference number: S57374; MUID:96021609; PMID:8533473
A:Accession: S57380
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1030 <ZUM>
A:Cross-references: EMBL:X83121; NID:9600461; PIDN:CA58190.1; PID:9600469
R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
submitted to the Protein Sequence Database, July 1996.
A:Reference number: S66775
A:Accession: S66783
A:Molecule type: DNA
A:Residues: 1-1030 <ZUM>
A:Cross-references: EMBL:Z74831; NID:91419931; PIDN:CA99101.1; PID:e251886; PID:9141993
A:Experimental source: strain S288C
C:Genetics:

A:Gene: SCD:HAL9
A:Cross-references: SGD:S0005449
A:Map position: 15L
C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluster
C:Keywords: transmembrane protein
F:131-171/Domain: GAL4 zinc binuclear cluster homology <GAL4>
F:762-778/Domain: transmembrane #status predicted <TM>

Query Match 71.2%; Score 37; DB 2; Length 1030;
Best Local Similarity 55.6%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SETINCHFK 9
Db 689 SRVLNCHFR 697

RESULT 5

probable finger protein YKR064w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Apr-2002
C:Accession: S38140
R:van Vliet-Reedijk, J.C.; Planta, R.J.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38130
A:Accession: S38140
A:Molecule type: DNA
A:Residues: 1-863 <VAN>
A:Cross-references: EMBL:Z28289; NID:9486530; PIDN:CA82143.1; PID:9486531; MIPS:YKR0
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0001772
A:Map position: 11R
C:Superfamily: unassigned GAL4-type zinc cluster proteins; GAL4 zinc binuclear cluste
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger
F:15-52/Domain: GAL4 zinc binuclear cluster homology <GAL4>
F:19-47/Region: zinc finger

Query Match 69.2%; Score 36; DB 2; Length 863;
Best Local Similarity 62.5%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ETINCHFK 9
Db 470 KTLQCHFK 477

RESULT 6

hypothetical protein H1406 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C:Accession: B64028
R:Reischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirness, E.F.; Kerlavage
; Goodyne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman
; D.M.; Brindon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64028
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-118 <TGR>
A:Cross-references: GB:U32820; GB:IA2023; NID:91574231; PIDN:AC23056.1; PID:91574244
C:Superfamily: Haemophilus influenzae hypothetical protein H1406

Query Match 67.3%; Score 35; DB 2; Length 118;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SETINCHFK 9
Db 35 SETVSCHFK 43

RESULT 7

Ly-49G.2 antigen - mouse
C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 17-Mar-2000
C:Accession: I49054
R:Smith, H.R.; Karhofer, F.M.; Yokoyama, W.M.
J. Immunol. 153, 1068-1079, 1994
A:Title: Ly-49 multigene family expressed by IL-2-activated NK cells.
A:Reference number: I49049; MUID:94300068; PMID:8027540
A:Accession: I49054
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-203 <RES>
A:Cross-references: EMBL:U10095; NID:9533497; PIDN:AAA50223.1; PID:9533498
C:Superfamily: natural killer cell receptor P1; C-type lectin homology

Query Match 67.3%; Score 35; DB 2; Length 203;
Best Local Similarity 83.3%; Pred. NO. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETINCH 7
DB 80 ETLNCH 85

RESULT 8
I49053
Ly-49G.2 antigen - mouse
N:Alternate names: Ly-49E
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000
C:Accession: I49053; I49115
R:Smith, H.R.; Karhofer, F.M.; Yokoyama, W.M.
J. Immunol. 153, 1068-1079, 1994
A:Title: Ly-49 multigene family expressed by IL-2-activated NK cells.
A:Reference number: I49049; MUID:94300068; PMID:8027540
A:Accession: I49053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-267 <RES>
A:Cross-references: EMBL:U10094; NID:9533495; PIDN:AAA50222.1; PID:9533496
R:Brennan, J.; Mager, D.; Jefferys, W.; Takel, F.
J. Exp. Med. 180, 2287-2295, 1994
A:Title: Expression of different members of the Ly-49 gene family defines distinct natu
A:Reference number: I49114; MUID:95053763; PMID:7964501
A:Accession: I49115
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-43, 'Q', 45-267 <RES>
A:Cross-references: EMBL:U12890; NID:9602409; PIDN:AAA58705.1; PID:9602410
C:Genetics:
A:Gene: Ly49G.4
C:Superfamily: natural killer cell receptor P1; C-type lectin homology

Query Match 67.3%; Score 35; DB 2; Length 267;
Best Local Similarity 83.3%; Pred. NO. 50;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETINCH 7
DB 80 ETLNCH 85

RESULT 9
I55686
LGL-1 - mouse
C:Species: Mus sp. (mouse)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 19-May-2000
C:Accession: I55686
R:Mason, L.H.; Orcaido, J.R.; Young, H.A.; Kumar, V.; Bennett, M.; Anderson, S.K.
J. Exp. Med. 182, 293-303, 1995
A:Title: Cloning and functional characteristics of murine large granular lymphocyte-1: a
A:Reference number: I55686; MUID:9535829; PMID:7629495
A:Accession: I55686
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-267 <RES>
A:Cross-references: GB:S78689; NID:91041888; PIDN:AAC60707.1; PID:91041889
C:Genetics:
A:Gene: LGL-1
C:Superfamily: natural killer cell receptor P1; C-type lectin homology

Query Match 67.3%; Score 35; DB 2; Length 267;
Best Local Similarity 83.3%; Pred. NO. 50;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETINCH 7
DB 80 ETLNCH 85

RESULT 10
I49052
Ly-49G.1 antigen - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C:Accession: I49052
R:Smith, H.R.; Karhofer, F.M.; Yokoyama, W.M.
J. Immunol. 153, 1068-1079, 1994
A:Title: Ly-49 multigene family expressed by IL-2-activated NK cells.
A:Reference number: I49049; MUID:94300068; PMID:8027540
A:Accession: I49052
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-280 <RES>
A:Cross-references: EMBL:U10093; NID:9533493; PIDN:AAA50221.1; PID:9533494
C:Superfamily: natural killer cell receptor P1; C-type lectin homology

Query Match 67.3%; Score 35; DB 2; Length 280;
Best Local Similarity 83.3%; Pred. NO. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETINCH 7
DB 80 ETLNCH 85

RESULT 11
D84968
Hypothetical protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: D84968
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: D84968
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-280 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: yjfr; BU328

Query Match 67.3%; Score 35; DB 2; Length 280;
Best Local Similarity 75.0%; Pred. NO. 52;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCH 8
DB 247 SRFINCH 254

RESULT 12
F95058
Hypothetical protein SP0505 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: F95058
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: F95058
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1426 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK74663.1; PID:914971978; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP0505

Query Match 67.3%; Score 35; DB 2; Length 426;
 Best Local Similarity 75.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCH 8
 DB 356 SGVINCH 363

RESULT 13

H97927
 type 1 site-specific deoxyribonuclease (EC 3.1.21.3) - Streptococcus pneumoniae (strain
 C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 27-Nov-2001
 C:Accession: H97927
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; Dehoff, B.S.; E
 y, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; Mcahren, S.; M
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: H97927
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1522 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK9252.1; PID:915458016; GSPDB:GN00174
 C:Genetics:
 A:Gene: hsdS
 C:Keywords: hydrolase

Query Match 67.3%; Score 35; DB 2; Length 522;
 Best Local Similarity 75.0%; Pred. No. 89;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCH 8
 DB 452 SGVINCH 459

RESULT 14

VCLJG2
 env polypeptide precursor - human immunodeficiency virus type 2 (isolate ROD)
 N:Alternate names: coat polypeptide
 C:Species: human immunodeficiency virus type 2, HIV-2
 C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C:Accession: C26262
 R:Guyader, M.; Emerman, M.; Sonigo, P.; Clavel, F.; Montagnier, L.; Alizon, M.
 Nature 326, 662-669, 1987
 A:Title: Genome organization and transactivation of the human immuno-deficiency virus ty
 A:Reference number: A26262; MUID:87173056; PMID:3031510
 A:Contents: proviral DNA
 A:Accession: C26262
 A:Molecule type: DNA
 A:Residues: 1-858 <GUY>

A:Cross-references: GB:M15390; NID:g1332361; PIDN:AA00770.1; PID:g325749
 C:Genetics:
 A:Gene: env
 C:Superfamily: type E retrovirus env polypeptide
 C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; poly
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-858/Product: env polypeptide #status predicted <NAV>
 F:18-501/Product: exterior membrane glycoprotein #status predicted <EXT>
 F:502-858/Product: transmembrane glycoprotein #status predicted <TM>
 F:34,67,76,119,120,151,166,179,192,193,196,206,238,241,248,272,278,289,300,367,371,40

Query Match 67.3%; Score 35; DB 1; Length 858;
 Best Local Similarity 85.7%; Pred. No. 14e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCH 8
 DB 159 ETINCH 165

RESULT 15

T25539
 hypothetical protein C10H11.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
 C:Accession: T25539
 R:Dante, M.; Wamsley, P.
 submitted to the EMBL data library, February 1997
 A:Description: The sequence of C. elegans cosmid C10H11.
 A:Reference number: Z20047
 A:Accession: T25539
 A:Status: preliminary; translated from GB/EMBL/DDB
 A:Molecule type: DNA
 A:Residues: 1-1173 <DAN>
 A:Cross-references: EMBL:U08311; PIDN:AA042348.1; GSPDB:GN00019; CESP:C10H11.9
 A:Experimental source: strain Bristol N2; clone C10H11
 C:Genetics:
 A:Gene: CESP:C10H11.9
 A:Map position: 1
 A:Introns: 23/3; 51/1; 104/3; 343/1; 478/2; 868/3; 891/3; 970/3; 1027/3; 1114/1
 C:Superfamily: hypothetical protein C10H11.9; protein kinase homology

Query Match 67.3%; Score 35; DB 2; Length 1173;
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 ETINCH 9
 DB 1091 ECKNCH 1098

Search completed: July 11, 2003, 12:00:31
 Job time : 2.91497 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 0.860707 seconds

(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-1

Perfect score: 52

Sequence: 1 SETINCHK 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	380	SPB7_HUMAN	O75635 homo sapien
2	37	71.2	320	AA3R_RAT	P28647 rattus norv
3	36	69.2	863	YK44_YEAST	P36023 saccharomyc
4	35	67.3	118	YE06_HAEM	P44181 haemophilus
5	35	67.3	280	KLR7_MOUSE	O60654 mus musculu
6	35	67.3	338	RSWC_BUCAL	P57413 buchnera ap
7	35	67.3	858	ENV_HV2RO	P04577 human immun
8	35	67.3	2095	RRLP_TOSV	P37800 toscana vir
9	35	67.3	2156	RPL_HUMAN	P56715 homo sapien
10	34	65.4	215	HB27_TAMSI	O06577 tamias sibi
11	34	65.4	231	HB2L_CHICK	P23068 gallus gall
12	34	65.4	436	ACHX_ONCYO	P54247 onchocerca
13	34	65.4	1078	CASR_HUMAN	P44180 homo sapien
14	34	65.4	1079	CASR_MOUSE	O94966 mus musculu
15	34	65.4	1079	CASR_RAT	P48442 rattus norv
16	34	65.4	1085	CASR_BOVIN	P35384 bos taurus
17	33	63.5	313	PENA_BURCE	O02940 burkholderi
18	33	63.5	319	AA3R_MOUSE	O64618 mus musculu
19	33	63.5	320	RLPA_RICPR	O94261 rickettsia
20	33	63.5	701	TREPE_XENLA	P02033 xenopus lae
21	33	63.5	764	PICR_HUMAN	P10183 homo sapien
22	32	61.5	69	CARP_SCHMA	P13566 schistosoma
23	32	61.5	92	NL7B_RICCO	P10974 ricinus com
24	32	61.5	178	RL25_HELPD	O921C4 heliobacte
25	32	61.5	226	14E_HUMAN	O14530 homo sapien
26	32	61.5	236	PLT_SHEEP	P16038 ovis aries
27	32	61.5	250	D250_ASFB7	P32092 african swi
28	32	61.5	277	BHWT_PIG	O95332 sus scrofa
29	32	61.5	287	SCR_CHICK	O09108 gallus gall
30	32	61.5	306	Y361_HAEM	P44662 haemophilus
31	32	61.5	320	RLPA_RICCN	O92183 rickettsia
32	32	61.5	327	ANX8_HUMAN	P13928 homo sapien
33	32	61.5	368	VG59_HSVSA	O01052 herpesvirus

34	32	61.5	406	1	BHWT_HUMAN	O93088 homo sapien
35	32	61.5	407	1	BHWT_MOUSE	O35490 mus musculu
36	32	61.5	407	1	BHWT_RAT	O09171 rattus norv
37	32	61.5	750	1	NU5C_TRECS	O91C2 tecoma scan
38	32	61.5	763	1	PEPX_LACLA	O9CE01 lactococcus
39	32	61.5	763	1	PEPX_LACLA	P22346 lactococcus
40	32	61.5	1254	1	POLS_EEYVA	P36331 venezuelan
41	32	61.5	1660	1	VIT6_OSCBR	O94637 oscobius br
42	32	61.5	3511	1	MY15_MOUSE	O94224 mus musculu
43	31.5	60.6	295	1	CIN5_YEAST	P40917 saccharomyc
44	31	59.6	130	1	LE62_MOUSE	O9C9W5 mus musculu
45	31	59.6	130	1	LE62_RAT	O92144 rattus norv

ALIGNMENTS

RESULT 1
SPB7_HUMAN STANDARD: PRT; 380 AA.
AC O75635;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Megsin (TP55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid:9606;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE-97326116; PubMed-9182567;
RA Tsumimoto M., Tsumoto K., Ishida N., Kurihara T., Iwasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.;
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity";
RT J. Biol. Chem. 272:15373-15380(1997).
RL [2].
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE-Mesangial cells;
RX MEDLINE-98376492; PubMed-9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, megslin, is a new serpin upregulated in
RT Iga nephropathy";
RT J. Clin. Invest. 102:828-836(1998).
RL [2].
RP FUNCTION: Might function as an inhibitor of Lys-specific
proteases. Might influence the maturation of megakaryocytes via
its action as a serpin.
CC [1] SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC [1] TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC [1] SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
or send an email to license@sdb.ch).
CC -----
CC EMBL: D88575; BAA1232.1; -
CC EMBL: AF027866; AAC64506.1; -
CC HSSP: P05619; 1HLE.
CC Genew: HGNC:13902; SERPINB7.
CC MIM: 603357; -
CC InterPro: IPR000215; Serpin.
CC Pfam: PF00079; serpin; 1.
CC SMART: SM00093; SERPIN; 1.
CC PROSITE: PS00284; SERPIN; 1.

KW Serpin: Serine protease inhibitor.
 FT ACT_SITE 347 348 REACTIVE_BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C3CFF605 CRC64;
 Query Match 100.0%; Score 52; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SETINCHFR 9
 Db 188 SETINCHFR 196

RESULT 2
 ID AA3R_RAT STANDARD; PRT; 320 AA.
 AC P28647; Q63792;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Adenosine A3 receptor (TGPCR1).
 GN ADORA3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE-Brain;
 RX MEDLINE=92366475; PubMed=1323836;
 RA Zhou O.Y., Li C.Y., Olan M.E., Johnson R.A., Stiles G.L., Civeilli O.;
 RT "Molecular cloning and characterization of an adenosine receptor: the
 RT A3 adenosine receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7432-7436(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Wistar; TISSUE-Testis;
 RX MEDLINE=91285122; PubMed=1647979;
 RA Meyerhof W., Mueller-Brechlin R., Richter D.;
 RT "Molecular cloning of a novel putative G-protein coupled receptor
 RT expressed during rat spermiogenesis.";
 RL FEBS Lett. 284:155-160(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=96196578; PubMed=8612733;
 RA Sajjadi F.G., Boyle D.L., Domingo R.C., Firestein G.S.;
 RT "cDNA cloning and characterization of A31, an alternatively spliced
 RT rat A3 adenosine receptor variant.";
 RL FEBS Lett. 382:125-129(1996).
 CC -1- FUNCTION: RECEPTOR FOR ADENOSINE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYL CYCLASE.
 CC -1- POSSIBLE ROLE IN REPRODUCTION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: TESTIS, PARTICULARLY IN SPERMATOCYTES AND
 CC SPERMATIDS BUT NOT IN SPERMATOGONIA. LOW LEVELS IN THE BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING SPERMIOGENESIS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X59249; CAA41937.1; -
 CC EMBL; M94152; AAA40680.1; -
 CC EMBL; X93219; CAA63702.1; -
 CC PIR; S17177; S17177.

DR PIR; A46152; A46152.
 DR HSSP; P29274; 1MMH.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7cm.1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G-PROTEIN_REC_P1.1; 1.
 DR PROSITE: PS50262; G-PROTEIN_REC_P1.2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Lipoprotein; Palmitate; Alternative splicing.
 FT DOMAIN 1 16
 FT TRANSMEM 17 39
 FT DOMAIN 40 50
 FT TRANSMEM 51 74
 FT DOMAIN 75 86
 FT TRANSMEM 87 108
 FT DOMAIN 109 128
 FT TRANSMEM 129 150
 FT DOMAIN 151 179
 FT TRANSMEM 180 200
 FT DOMAIN 201 233
 FT TRANSMEM 234 257
 FT DOMAIN 258 263
 FT TRANSMEM 264 286
 FT DOMAIN 287 320
 FT CARBOHYD 4 4
 FT CARBOHYD 5 5
 FT DISULFID 85 168
 FT LIPID 305 305
 FT VARSPPLIC 119 119
 FT CONFLICT 18 18
 FT CONFLICT 68 68
 FT CONFLICT 74 83
 FT CONFLICT 132 132
 FT CONFLICT 172 172
 FT CONFLICT 288 288
 FT CONFLICT 291 295
 SQ SEQUENCE 320 AA; 36629 MW; 59DBA40C594ACDAB CRC64;
 Query Match 71.2%; Score 37; DB 1; Length 320;
 Best Local Similarity 55.6%; Pred. No. 5.3;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SETINCHFR 9
 Db 163 SSTLSCHFR 171

RESULT 3
 ID YK44_YEAST STANDARD; PRT; 863 AA.
 AC P36023;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative 101.8 kDa transcriptional regulatory protein in *YAS1-CCP1*
 DE intergenic region.
 GN YKR064W.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288c;
 RA van Vlier-Reedijk J.C., Planta R.J.;
 RU Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
 CC CLUSTER DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its


```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z28289; CAA82143.1;
DR PIR: S38140;
DR HSSP: P12351; 1HWT.
DR SGD: S0001772; YKR064W.
DR InterPro: IPR001138; FunG1-TrN.
DR Pfam: PF00172; Zn_c1us; 1.
DR SMART: SMO0066; GAL4; 1.
DR PROSITE: PS00463; ZN2-CY6-FUNGAL_1; 1.
DR PROSITE: PS50048; ZN2-CY6-FUNGAL_2; 1.
RW Hypothetical protein: Transcription regulation; DNA-binding;
FT Nucleic protein: zinc; Metal-binding.
FT DNM_BIND 18 47 ZN(2)-CYS(6), FUNGAL-TYPE.
SO SEQUENCE 863 AA; 101824 MW; B1A04655405FAFB4 CRC64;

Oy 2 ETINCHK 9
Db 470 KTLCHK 477

Query Match 69.2%; Score 36; DB 1; Length 863;
Best Local Similarity 62.5%; Pred. No. 21;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

ID YE06_HAEIN STANDARD: PRT; 118 AA.
YE06_HAEIN
AC P44181;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11406.
DE H11406.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL Rd."
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on lis-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32820; AAC23056.1;
DR TIGR: H11406;
RW Hypothetical protein; Complete proteome.
SO SEQUENCE 118 AA; 13325 MW; C1CAF33E6CB3452 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 118;
Best Local Similarity 66.7%; Pred. No. 5;

```

	Matches	6:	Conservative	2:	Mismatches	1:	Indels	0:	Gaps	0:
OY		1	SETINCHFK 9							
		:								
Db		35	SETVSYHK 43							
	RESULT 5									
	KLRT_MOUSE									
ID	KLRT_MOUSE	STANDARD;		PRT;	280	AA.				
AC	060654;	060655;	060656;	060683;						
DT	01-NOV-1997	(Rel. 35,	Created)							
DT	01-NOV-1997	(Rel. 35,	Last sequence update)							
DT	16-Oct-2001	(Rel. 40,	Last annotation update)							
DE	Killer cell lectin-like receptor 7 (T-cell surface glycoprotein									
DE	LY-49G) (LY49-G antigen).									
GN	KLRT7 OR LY49G OR LY-49G OR LY49-G OR LY49G4.									
OS	Mus musculus (Mouse).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
OX	NCBI_TaxID=10090;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=C57BL/6; TISSUE=Spleen;									
RX	MEDLINE=94300068; PubMed=8027540;									
RA	Smith H.R.C., Kalinhoer F.M., Yokoyama W.M.;									
RT	"Ly-49 murine gene family expressed by IL-2-activated NK cells.";									
RL	J. Immunol. 153:1068-1079(1994).									
RN	[2]									
RP	SEQUENCE FROM N.A. (ISOFORM LY-49G.2).									
RC	STRAIN=B10.A;									
RX	MEDLINE=95053763; PubMed=7964501;									
RA	Brennan J., Mager D., Jefferies W., Takel F.;									
RT	"Expression of different members of the Ly-49 gene family defines									
RL	distinct natural killer cell subsets and cell adhesion properties.";									
CC	J. Exp. Med. 180:2287-2295(1994).									
CC	-1- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR CLASS I MHC.									
CC	-1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.									
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein.									
CC	-1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: LY-49G.1 (SHOWN HERE), LY-									
CC	49G.2 AND LY-49G.3: ARE PRODUCED BY ALTERNATIVE SPLICING.									
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.									
CC	-----									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration									
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -									
CC	the European Bioinformatics Institute. There are no restrictions on its									
CC	use by non-profit institutions as long as its content is in no way									
CC	modified and this statement is not removed. Usage by and for commercial									
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/									
CC	or send an email to license@isb-sib.ch).									
CC	-----									
DR	EMBL: U10093; AAA50322.1; -									
DR	EMBL: U10094; AAA50322.1; -									
DR	EMBL: U10095; AAA50223.1; -									
DR	EMBL: U12890; AAA58705.1; -									
DR	MGI: 101901; Klrp7.									
DR	InterPro: IPR001304; Lectin_C.									
DR	pfam: PF00059; lectin_c.1.									
DR	SMART: SM00034; CLECT.1.									
DR	PROSITE: PS00615; C-TYPE_LECTIN_1; FALSE_NEG.									
DR	PROSITE: PS50041; C-TYPE_LECTIN_2; 1.									
KW	T-cell; glycoprotein; Antigen; Transmembrane; Cell adhesion;									
KW	Signal-anchor; Lectin; Receptor; Multigene family;									
KW	Alternative splicing.									
FT	DOMAIN	1	44							

```

FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VANSPLIC 117 193 MISSING (IN ISOFORM LY-49G3).
FT VANSPLIC 142 154 MISSING (IN ISOFORM LY-49G2).
FT CONFLICT 44 44 K -> O (IN REF. 2).
SQ SEQUENCE 280 AA; 32522 MW; 4704A2D87D5E83FB CRC64;

Query Match 67.3%; Score 35; DB 1; Length 280;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETINCH 7
Db 80 ETINCH 85

RESULT 6
RSMC_BUCAL STANDARD; PRT; 338 AA.
ID RSMC_BUCAL
AC P57413;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosomal RNA small subunit methyltransferase C (EC 2.1.1.52) (rRNA
DE (guanine-N(2)-methyltransferase) (16S rRNA m2G1207
DE methyltransferase).
GN RSMC OR B0328.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Tokyo 1998; PubMed=10993077;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. Aps.";
RL Nature 407:81-86(2000).
CC -1- FUNCTION: SPECIFICALLY METHYLATES THE GUANOSINE IN POSITION 1207
CC OF 16S rRNA IN THE 30S PARTICLE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + RNA = S-adenosyl-L-
CC homocysteine + RNA containing N(2)-methylguanine.
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. RSMC
CC SUPERFAMILY.
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A POTENTIAL FRAMESHIFT
CC WAS CORRECTED IN POSITION 57.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AP001119; BAB13036.1; ALT_FRAME.
DR InterPro: IPR002052; N6_Mtase.
DR InterPro: IPR000051; SAM_Bind.
DR PROSITE: PS00092; N6_Mtase; UNKNOWN_1.
KM rRNA processing; Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 338 AA; 39105 MW; E43516DD22FA014 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 338;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SETINCH 8
Db 305 SREINCH 312

RESULT 7
ENV_HVZRO

```

```

ID ENV_HVZRO STANDARD; PRT; 858 AA.
AC P04577;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope glycoprotein gp160 precursor [contains: Exterior membrane
DE glycoprotein gp160; Transmembrane glycoprotein (GP41)].
DE ENV.
GN Envelope glycoprotein gp160 precursor [contains: Exterior membrane
GN glycoprotein gp160; Transmembrane glycoprotein (GP41)].
OS Human immunodeficiency virus type 2 (isolate ROD) (HIV-2).
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11720;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87173056; PubMed=3031510;
RA Guyader M., Emerman M., Sonigo P., Clavel F., Montagnier L.,
RA Alizon M.;
RT "Genome organization and transactivation of the human
RT immunodeficiency virus type 2.";
RL Nature 326:662-669(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M15390; AAB00770.1; -
DR EMBL, X05291; CAA28914.1; -
DR PIR, C26282; VCLD2.
DR HIV, M15390; ENV52ROD.
DR InterPro: IPR000328; ENV_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam, PF00516; GP120.1.
DR Pfam, PF00517; GP41.1.
DR AIDS; Coat protein; Glycoprotein; Transmembrane;
DR Signal.
KW CHAIN
FT CHAIN 18 501
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 312 312 T -> I (IN REF. 1; AAB00770).
SQ SEQUENCE 858 AA; 98824 MW; C7266AFLF5C5B9A7 CRC64;

```

Query Match 67.3%; Score 35; DB 1; Length 858;
 Best Local Similarity 85.7%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCHP 8
 DB 159 ETINCOF 165

RESULT 8
 ID RPL_TOSV STANDARD; PRT: 2095 AA.
 AC P37800;
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
 GN L.
 OS Toscana virus (Tos).
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
 OX NCBI_TaxID=11590;
 RN 11
 RC SEQUENCE FROM N.A.
 RC STRAIN=ISS.PHL.3;
 RX MEDLINE=9312587; PubMed=8460526;
 RA Accardi L., Gro M.C., di Bonito P., Giorgi C.;
 RT "Toscana virus genomic L segment: molecular cloning, coding strategy
 RT and amino acid sequence in comparison with other negative strand RNA
 RT viruses.";
 RL Virus Res. 27:119-131(1993).
 CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X68414; CAA48478.1; -
 CC DR PIR: S29529; S29529.
 CC KM Transferrase: Nucleotidyltransferase; RNA-directed RNA polymerase.
 CC SO SEQUENCE 2095 AA; 238885 MW; 22FF4DAD745583B7 CRC64;
 CC -----
 CC Query Match 67.3%; Score 35; DB 1; Length 2095;
 CC Best Local Similarity 62.5%; Pred. No. 74;
 CC Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCHP 9
 DB 501 ETLDCCHK 508

RESULT 9
 ID RPL_HUMAN STANDARD; PRT: 2156 AA.
 AC P56715;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oxygen-regulated protein 1 (Retinitis pigmentosa RPL protein)
 DE (Retinitis pigmentosa 1 protein).
 GN RPL OR ORP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A., AND VARIANTS HIS-872; TYR-985; THR-1670; PRO-1691
 RP AND TYR-2033.

RC TISSUE=Retina;
 RX MEDLINE=99318096; PubMed=10391212;
 RA Sullivan L.S., Heckenlively J.R., Byrne S.J., Zuo J., Hyde M.A.,
 RA Gal A., Denton M., Inglehearn C.F., Blanton S.H., Dalgier S.P.,
 RT "Mutations in a novel retina-specific gene cause autosomal dominant
 RT retinitis pigmentosa.";
 RL Nat. Genet. 22:255-259(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=99318095; PubMed=10391211;
 RA Pierce E.A., Quinn T., Mehan T., McGee T.L., Berson E.L., Dryja T.P.,
 RT "Mutations in a gene encoding a new oxygen-regulated photoreceptor
 RT protein cause dominant retinitis pigmentosa.";
 RL Nat. Genet. 22:248-254(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99330563; PubMed=10401003;
 RA Guillonneau X., Parfey N.I., Danciger M., Kozak C.A., Clidecyan A.V.,
 RA Jacobson S.G., Farber D.B.,
 RT "A nonsense mutation in a novel gene is associated with retinitis
 RT pigmentosa in a family linked to the RPL locus.";
 RL Hum. Mol. Genet. 8:1541-1546(1999).
 CC -1 FUNCTION: COULD HAVE A ROLE IN THE DIFFERENTIATION OF
 CC PHOTORECEPTOR CELLS.
 CC -1 TISSUE SPECIFICITY: EXPRESSED IN RETINA. NOT EXPRESSED IN HEART,
 CC BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, SPLEEN AND
 CC PANCREAS.
 CC -1 DISEASE: DEFECTS IN RPL CAUSE RETINITIS PIGMENTOSA FORM 1 (RPL); A
 CC DISEASE CHARACTERIZED BY CONSTRUCTION OF THE VISUAL FIELDS, NIGHT
 CC BLINDNESS, AND FUNDUS CHANGES. THE DISEASE SEEMS TO BE ASSOCIATED
 CC WITH TRUNCATED (STOP OR FRAMESHIFT MUTATIONS) FORMS OF THE
 CC PROTEIN.
 CC -1 SIMILARITY: CONTAINS 2 DOUBLECORTIN DOMAINS.
 CC -1 DATABASE: NAME=Retnet;
 CC NOTE=Retinal information network;
 CC WWW="http://www.sph.utmc.edu/retnet/".
 CC -1 DATABASE: NAME=Mutations of the RPL gene;
 CC NOTE=Retina International's Scientific Newsletter;
 CC WWW="http://www.retina-international.com/sci-news/rplmut.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AF143226; AAD44197.1; -
 CC DR EMBL: AF143224; AAD44197.1; JOINED.
 CC DR EMBL: AF143225; AAD44197.1; JOINED.
 CC DR EMBL: AF143222; AAD44198.1; -
 CC DR EMBL: AF141021; AAD42072.1; -
 CC DR EMBL: AF152242; AAD46774.1; -
 CC DR EMBL: AF152240; AAD46774.1; JOINED.
 CC DR EMBL: AF152241; AAD46774.1; JOINED.
 CC DR EMBL: AF14592; AAD46769.1; -
 CC DR Genew; HGNC:10263; RPL.
 CC DR MIM: 603937; -
 CC DR MIM: 180100; -
 CC DR InterPro: IPR003533; DCX.
 CC DR Pfam: PF03607; DCX; 2.
 CC DR SMART: SM00537; DCX; 2.
 CC DR PROSITE: PS50309; DCX; 2.
 CC DR Vision; Retinitis pigmentosa; Polymorphism; Repeat.
 CC FT DOMAIN 36 118 DOUBLECORTIN 1.
 CC FT FT 154 233 DOUBLECORTIN 2.
 CC FT DOMAIN 268 273 POLY-SER.
 CC FT DOMAIN 671 675 POLY-LYS.
 CC FT DOMAIN 1687 1691 POLY-SER.
 CC FT VARIANT 872 872 R -> H.
 CC /FTID=VAR_007810.

FT VARIANT 985 985 N->Y. /FTID=VAR_007811.
 FT VARIANT 1670 1670 A->T. /FTID=VAR_007812.
 FT VARIANT 1691 1691 S->P. /FTID=VAR_007813.
 FT VARIANT 2033 2033 C->Y. /FTID=VAR_007814.
 FT SEQUENCE 2156 AA; 240659 MW; 55AEDEBC43D6A507 CRC64;
 Query Match 67.3%; Score 35; DB 1; Length 2156;
 Best Local Similarity 71.4%; Pred. No. 76;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCHP 8
 Db 2089 ENINCHY 2095

RESULT 10
 HP27_TAMSI STANDARD; PRT; 215 AA.
 AC 006577;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hibernation-associated plasma protein HP-27 precursor (Hibernator-specific blood complex, 27 kDa subunit).
 OS Tamias sibiricus (Siberian chipmunk) (Asian chipmunk).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
 OC Tamias.
 OX NCBI_TaxID=64680;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver.
 RX MEDLINE=93180798; PubMed=8441393.
 RA Takamatsu N., Oba K., Kondo J., Kondo N., Shiba T.;
 RT "Hibernation-associated gene regulation of plasma proteins with a collagen-like domain in mammalian hibernators.";
 RL Mol. Cell. Biol. 13:1516-1524(1993).
 RN [2]
 RP SEQUENCE OF 31-215.
 RC TISSUE=Plasma;
 RX MEDLINE=92112696; PubMed=1730610;
 RA Kondo N., Kondo J.;
 RT "Identification of novel blood proteins specific for mammalian hibernation.";
 RL J. Biol. Chem. 267:473-478(1992).
 CC -1- FUNCTION: PLASMA PROTEINS HP-20, HP-25, HP-27 AND HP-55 FORM A 140 KDA COMPLEX VIA DISULFIDE BONDS IN THE PLASMA AND ARE HIBERNATION SPECIFIC.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
 CC -1- DEVELOPMENTAL STAGE: THE PROTEIN COMPLEX DISAPPEARS FROM THE PLASMA AT ONSET OF HIBERNATION AND REAPPEARS AS HIBERNATION CEASES.
 CC -1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C10 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D12976; BAA02353.1; -
 DR Interpro: IPR001073; C1Q.
 DR Interpro: IPR000087; Collagen.
 DR Pfam: PF00386; C1q; 1.
 DR Pfam: PF01391; Collagen; 1.

DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR SMART; SM00110; C1Q; 1.
 DR PROSITE; PS01113; C1Q; 1.
 KW Signal; collagen; glycoprotein; Plasma; Multigene family.
 FT SIGNAL 1 30
 FT CHAIN 31 215
 FT DOMAIN 43 81
 FT DOMAIN 83 215
 FT CARBOHYD 155 155
 FT CONFLICT 41 41
 FT CONFLICT 51 51
 FT SEQUENCE 215 AA; 22797 MW; D2357086A08FA9E CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 215;
 Best Local Similarity 71.4%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 TINCHP 9
 Db 81 TVNCHSK 87

RESULT 11
 HB2L_CHICK STANDARD; PRT; 231 AA.
 AC P23068;
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-NOV-1991 (Rel. 20, Last annotation update)
 DE Class II histocompatibility antigen, B-L beta chain (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A. (CLONE P14).
 RX MEDLINE=88296406; PubMed=2841107;
 RA Bourlet Y., Behar G., Guillemot F., Frechin N., Billault A., Chausse A.M., Zoorob R., Auffray C.;
 RT "Isolation of chicken major histocompatibility complex class II (B-L) beta chain sequences: comparison with mammalian beta chains and expression in lymphoid organs.";
 RL EMBO J. 7:1031-1039(1988).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X07447; -; NOT_ANNOTATED_CDS.
 DR PIR; S00475; HLCBBL.
 DR HSSP; P01888; IBMG.
 DR Interpro: IPR003006; Ig_MHC.
 DR Interpro: IPR003597; Ig_c1.
 DR Interpro: IPR000353; MHC_II_beta.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR Prodom; PD000328; MHC_II_beta; 1.
 DR SMART; SM00407; IgC1; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW MHC II; Transmembrane; Glycoprotein.
 KW NON_TER 1
 FT DOMAIN 1 89
 FT DOMAIN 90 182
 FT DOMAIN 183 194
 FT TRANSMEM 195 219
 FT DOMAIN 220 231
 FT DISULFID 10 74
 CC EXTRACELLULAR BETA-1.
 CC EXTRACELLULAR BETA-2.
 CC CONNECTING PEPTIDE.
 CC CYTOPLASMIC TAIL.
 CC BY SIMILARITY.

FT DISULFID 111 167 BY SIMILARITY.
 FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 231 AA; 26278 MW; 06ADE95FAFB222A CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 231;
 Best Local Similarity 62.5%; Pred. No. 15;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SETINCHP 8
 DB 5 SATVECHP 12
 RESULT 12
 ACHX_ONCVO
 ID ACHX_ONCVO STANDARD; PRT; 436 AA.
 AC P34247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetylcholine receptor protein, non-alpha chain (Fragment).
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioida;
 OC Onchocercidae; Onchocerca.
 OX NCBI_TaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94299155; PubMed=8026747;
 RA Ajuu P.M., Eswang T.G.;
 RT "Cloning of a cDNA encoding a putative nicotinic acetylcholine
 receptor subunit of the human filarial parasite Onchocerca
 volvulus.";
 RT Gene 144:127-129(1994).
 CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
 EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
 LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
 MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sib.ch).
 CC -----
 DR EMBL: L20465; AAA21823.1; .
 DR EMBL: L12543; AAA29415.1; .
 DR InterPro: IPR000188; GABA_A_receptor.
 DR InterPro: IPR001175; Neur_Channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR ProSite: PS00236; NEUROTRION_CHANNEL; 1.
 KM Postsynaptic membrane; Ionic channel; Glycoprotein; Transmembrane;
 KM Multigene family.
 FT NON_TER 1
 FT DOMAIN 1
 FT TRANSMEM 195
 FT TRANSMEM 196
 FT TRANSMEM 219
 FT TRANSMEM 227
 FT TRANSMEM 245
 FT DOMAIN 261
 FT DOMAIN 280
 FT TRANSMEM 281
 FT TRANSMEM 404
 FT TRANSMEM 405
 FT TRANSMEM 423
 FT DISULFID 89
 FT CARBOHYD 103
 FT CARBOHYD 62
 FT CARBOHYD 62
 FT CARBOHYD 140
 FT CARBOHYD 140
 SO SEQUENCE 436 AA; 51340 MW; 58051BFES9A4555F CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 436;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 NCHFR 9
 DB 102 NCHFR 106
 RESULT 13
 CASR_HUMAN
 ID CASR_HUMAN STANDARD; PRT; 1078 AA.
 AC P41180; Q13912; Q16379; Q16108; Q16109; Q16110;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid
 DE cell calcium-sensing receptor).
 GN CASR OR GPRC2A OR PCARL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Pearce S.H.S., Thakker R.V.;
 RT Submitted (Dec-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Tissue-Parathyroid;
 RX MEDLINE=95279439; PubMed=7759551;
 RA Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,
 RA Hebert S.C., Nemeth E.F., Fuller F.;
 RT "Molecular cloning and functional expression of human parathyroid
 RT calcium receptor cDNAs.";
 RT J. Biol. Chem. 270:12919-12925(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC Tissue-Kidney;
 RX MEDLINE=95408281; PubMed=7677761;
 RA Aida K., Koishi S., Tawata M., Onaya T.;
 RT "Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from
 RT human kidney.";
 RT Biochem. Biophys. Res. Commun. 214:524-529(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96343808; PubMed=8756555;
 RA Freichel M., Zink-Lorenz A., Holltschl A., Hafner M., Flockerzi V.,
 RA Raue F.;
 RT "Expression of a calcium-sensing receptor in a human medullary
 RT thyroid carcinoma cell line and its contribution to calcitonin
 RT secretion.";
 RT J. Clin. Invest. 97:1085-1093(1996).
 RN [5]
 RP SEQUENCE OF 643-908 FROM N.A.
 RX MEDLINE=96193893; PubMed=8613532;
 RA Bikle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;
 RT "Changes in calcium responsiveness and handling during keratinocyte
 RT differentiation. Potential role of the calcium receptor.";
 RT J. Clin. Invest. 97:1085-1093(1996).
 RN [6]
 RP VARIANTS FHH GLN-185; LYS-297 AND TRP-795.
 RX MEDLINE=94094324; PubMed=7916660;
 RA Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,
 RA Steinmann B., Levy T., Seidman C.E., Seidman J.C.;
 RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
 RT hypocalcaemic hypercalcaemia and neonatal severe
 RT hyperparathyroidism.";
 RT Cell 75:1297-1303(1993).
 RN [7]
 RP VARIANT ADH ALA-127.
 RX MEDLINE=95179179; PubMed=7874174;
 RA Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Kifor O., Park J.,
 RA Hebert S.C., Seidman C.E., Seidman J.C.;
 RT "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
 RT gene mutation.";
 RT Nat. Genet. 8:303-307(1994).

[8]
 RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
 RX MEDLINE-95243222; PubMed-7726161;
 RA Chou Y.-H.W., Pollak M.R., Brandt M.L., Toss G., Arngvist H.,
 RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seldman J.G.,
 RA Seldman C.E.;
 RT "Mutations in the human Ca(2+)-sensing-receptor gene that cause
 RT familial hypocalcemic hypercalcaemia";
 RL Am. J. Hum. Genet. 56:1075-1079(1995).
 RN [9]
 RP SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
 RX MEDLINE-95403641; PubMed-7673400.
 RA Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T.;
 RT "Familial hypocalcemic hypercalcaemia associated with mutation in the
 RT human Ca(2+)-sensing receptor gene";
 RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
 RN [10]
 RP VARIANTS NSHPT LEU-227 AND TYR-598.
 RX MEDLINE-96292293; PubMed-8675635;
 RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
 RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,
 RA Thakker R.V.;
 RT "Calcium-sensing receptor mutations in familial benign hypercalcaemia
 RT and neonatal hyperparathyroidism";
 RL J. Clin. Invest. 96:2683-2692(1995).
 RN [11]
 RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
 RX MEDLINE-96311554; PubMed-8733126;
 RA Baron J., Miner K.K., Yanovski J.A., Cunningham A.W., Iane L.,
 RA Zimmerman D., Cutler G.B. Jr.;
 RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
 RT dominant and sporadic hyperparathyroidism";
 RL Hum. Mol. Genet. 5:601-606(1996).
 RN [12]
 RP VARIANT FHH ARG-174.
 RX MEDLINE-97442275; PubMed-9298824.
 RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laling N.G., Pullan P.T.,
 RA Ratajczak T.;
 RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene
 RT associated with familial hypocalcemic hypercalcaemia";
 RL Hum. Mutat. 10:233-235(1997).
 RN [13]
 RP VARIANT FHH GLU-557.
 RX MEDLINE-21603857; PubMed-11762699;
 RA Nakayama T., Minato M., Nakagawa M., Soma M., Tobe H., Aoi N.,
 RA Kosuge K., Sato M., Ozawa Y., Kanatsuse K., Kokubun S.;
 RT "A novel mutation in Ca2+-sensing receptor gene in familial
 RT hypocalcemic hypercalcaemia";
 RL Endocrine 15:277-282(2001).
 CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,
 CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
 CC -1- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC
 CC HYPERCALCAEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM
 CC (NSHPT). TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM
 CC HOMEOSTASIS. THE MUTATIONS REDUCE THE ACTIVITY OF THE RECEPTOR.
 CC FHH AFFECTED INDIVIDUALS EXHIBIT MILD OR MODEST HYPERCALCAEMIA,
 CC RELATIVE HYPOCALCAEMIA, AND INAPPROPRIATELY NORMAL PTH LEVELS. IN
 CC CONTRAST, NSHPT IS A RARE AUTOSOMAL RECESSIVE LIFE-THREATENING
 CC DISORDER CHARACTERIZED BY VERY HIGH SERUM CALCIUM CONCENTRATIONS,
 CC SKELETAL DEMINERALIZATION, AND PARATHYROID HYPERPLASIA. IN SOME
 CC INSTANCES NSHPT HAS BEEN DEMONSTRATED TO BE THE HOMOZYGOUS FORM OF
 CC FHH.
 CC -1- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
 CC HYPOCALCAEMIA (ADH) IN WHICH THE RECEPTOR IS ACTIVATED AT SUBNORMAL
 CC CA(2+) LEVELS.
 CC -1- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT

CC HYPOPARATHYROIDISM (ADHP). ADHP IS CHARACTERIZED BY HYPOCALCAEMIA
 CC AND HYPERPHOSPHATEMIA DUE TO INADEQUATE SECRETION OF PARATHYROID
 CC HORMONE. SYMPTOMS ARE SEIZURES, TETANY AND CRAMPS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X81086; CAAS6990.1; -
 CC DR EMBL: U20759; AAA86503.1; -
 CC DR EMBL: U20760; AAA86504.1; -
 CC DR EMBL: D50855; BAA09453.1; -
 CC DR EMBL: S83176; AAB46873.1; -
 CC DR EMBL: S79217; AAB35262.2; -
 CC DR EMBL: S68032; AAB29413.2; ALT_SEQ.
 CC DR EMBL: S68033; AAB29414.1; -
 CC DR EMBL: S68036; AAB29415.1; -
 CC DR EMBL: S81735; AAD14370.1; -
 CC DR Genew: HGNC:1514; CASR.
 CC DR MIM: 601199; -
 CC DR MIM: 145980; -
 CC DR MIM: 239200; -
 CC DR MIM: 601198; -
 CC DR InterPro: IPR001828; ANF_receptor.
 CC DR InterPro: IPR000337; GPCR_Mgt.
 CC DR Pfam: PF00003; 7tm_3; 1.
 CC DR Pfam: PF01094; ANF_receptor; 1.
 CC DR PRINTS: PR00248; GPCRMR.
 CC DR PROSITE: PS00979; G-PROTEIN_REC_P3_1; 1.
 CC DR PROSITE: PS00980; G-PROTEIN_REC_P3_2; 1.
 CC DR PROSITE: PS00981; G-PROTEIN_REC_P3_3; 1.
 CC DR PROSITE: PS02593; G-PROTEIN_REC_P3_4; 1.
 CC DR PROSITE: PS02593; G-PROTEIN_REC_P3_4; 1.
 CC DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 CC KW Disease mutation; Alternative splicing; Polymorphism.
 CC FT SIGNAL 1 19
 CC FT CHAIN 20 1078
 CC FT DOMAIN 20 612
 CC FT TRANSSEM 613 635
 CC FT DOMAIN 636 649
 CC FT TRANSSEM 650 670
 CC FT DOMAIN 671 681
 CC FT TRANSSEM 682 700
 CC FT DOMAIN 701 724
 CC FT TRANSSEM 725 745
 CC FT DOMAIN 746 769
 CC FT TRANSSEM 770 792
 CC FT DOMAIN 793 805
 CC FT TRANSSEM 806 828
 CC FT DOMAIN 829 836
 CC FT TRANSSEM 837 862
 CC FT DOMAIN 863 1078
 CC FT CARBOHYD 90 130
 CC FT CARBOHYD 130 130
 CC FT CARBOHYD 261 261
 CC FT CARBOHYD 287 287
 CC FT CARBOHYD 386 386
 CC Query Match. 65.4%; Score 34; DB 1; Length 1078;
 CC Best Local Similarity 62.5%; Pred. No. 62;
 CC Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 CC QY 2 ETINCHER 9
 CC II III;
 CC Db 354 ETFNCHLQ 361
 CC RESULT 14
 CC CASR_MOUSE

ID CASR_MOUSE STANDARD: PRT: 1079 AA.
 AC 090Y96; 008968; 088519; 090Y95; 090Z08; 09R1D6; 09R1Y2;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPRC2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID:10090;
 [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC STRAIN-C57BL/6; TISSUE-Kidney;
 RX MEDLINE-20092890; PubMed-10625662;
 RA Oda Y., Tu C.-L., Chang W., Crumrine D., Koenigues L., Mauro T.,
 RA Elias P.M., Bkile D.D.;
 RT "The calcium sensing receptor and its alternatively spliced form in
 RT murine epidermal differentiation.";
 RL J. Biol. Chem. 275:1183-1190(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC STRAIN-Black Swiss X 129/SVJ; TISSUE-Kidney;
 RX MEDLINE-20119279; PubMed-10652312;
 RA P.I., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
 RT "Sensing of extracellular cations in Casr-deficient osteoblasts.
 RT Evidence for a novel cation-sensing mechanism.";
 RL J. Biol. Chem. 275:3256-3263(2000).
 RN [3]
 RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC TISSUE-Epiphyseal cartilage;
 RX MEDLINE-20043955; PubMed-10579354;
 RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
 RA Miller S., Shoback D.;
 RT "Expression and signal transduction of calcium-sensing receptors in
 RT cartilage and bone.";
 RL Endocrinology 140:5883-5893(1999).
 RN [4]
 RP SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
 RC STRAIN-NMRI; TISSUE-Brain;
 RA Hildengrand J., Ammon H.P.T., Wahl M.A.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 562-814 FROM N.A.
 RC TISSUE-Kidney;
 RA Moswad T.T., Riccardi D.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE-97231187; PubMed-9076582;
 RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hinson T.K.;
 RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
 RT functionally related to the calcium receptor.";
 RL J. Bone Miner. Res. 12:393-402(1997).
 CC -1- FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EPIDERMIS, KIDNEY AND CARTILAGE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AF110178; AAD28371.1; -
 DR EMBL: AF110179; AAD28372.1; -
 DR EMBL: AF128842; AAD40638.1; -
 DR EMBL: AF068900; AAC19388.1; -
 DR EMBL: AB027140; BAA77688.1; -
 DR EMBL: AF002015; AAC53252.1; -
 DR EMBL: AF159565; AAF00193.1; -
 DR MGI: 1351351; Gprc2a.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_REC_P3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_REC_P3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_P3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_REC_P3_4; 1.
 DR G-protein coupled receptor; Transmembrane; glycoprotein; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 19
 FT CHAIN 20 1079
 FT DOMAIN 20 612
 FT TRANSMEM 613 635
 FT DOMAIN 636 649
 FT TRANSMEM 650 670
 FT DOMAIN 671 681
 FT TRANSMEM 682 700
 FT DOMAIN 701 724
 FT TRANSMEM 725 745
 FT DOMAIN 746 769
 FT TRANSMEM 770 792
 FT DOMAIN 793 805
 FT TRANSMEM 806 828
 FT TRANSMEM 829 836
 FT DOMAIN 837 862
 FT TRANSMEM 863 1079
 FT CARBOHYD 90 90
 FT CARBOHYD 130 130
 FT CARBOHYD 261 261
 FT CARBOHYD 287 287
 FT CARBOHYD 386 386
 FT CARBOHYD 446 446
 FT CARBOHYD 468 468
 FT CARBOHYD 488 488
 FT CARBOHYD 541 541
 FT CARBOHYD 594 594
 FT VARSPLIC 461 537
 FT CONFLICT 45 45
 FT CONFLICT 304 304
 FT CONFLICT 410 410
 FT CONFLICT 566 566
 FT CONFLICT 595 595
 FT CONFLICT 610 610
 FT CONFLICT 814 814
 FT CONFLICT 889 889
 FT CONFLICT 906 909
 FT CONFLICT 1057 1057
 FT CONFLICT 1064 1064
 FT CONFLICT 1076 1076
 SQ SEQUENCE 1079 AA; 120839 MW; AAF8B8DA72736D6E CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 1079;
 Best Local Similarity 62.5%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 2 EFINCHF 9
 Db 354 EFINCHLQ 361
 RESULT 15
 CASR_RAT
 ID CASR_RAT STANDARD: PRT: 1079 AA.

AC P48442;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPRC2A OR PCARL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Kidney outer medulla;
 RX MEDLINE=95116508; PubMed=7816802;
 RA Ricciardi D., Park J., Lee W., Gamba G., Brown E.M., Hebert S.C.;
 RT "Cloning and functional expression of a rat kidney extracellular
 RT calcium/polyvalent cation-sensing receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:131-135(1995).
 RN [2]
 RP SEQUENCE OF 1-294 FROM N.A.
 RC STRAIN-Wistar;
 RX MEDLINE=95241465; PubMed=7724534;
 RA Ruat M., Snokman A.M., Snyder S.H.;
 RT "Calcium sensing receptor: molecular cloning in rat and localization
 RT to nerve terminals.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3161-3165(1995).
 CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U10354; AAC52149.1; -;
 DR EMBL; U20289; AAC52195.1; -;
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCR_MGR.
 DR PROSITE; PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR PROSITE; PS50259; G_PROTEIN_RECPE_F3_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1
 FT CHAIN 20
 FT DOMAIN 20
 FT TRANSSEM 613
 FT DOMAIN 635
 FT TRANSSEM 636
 FT TRANSSEM 650
 FT TRANSSEM 670
 FT TRANSSEM 671
 FT TRANSSEM 681
 FT TRANSSEM 682
 FT TRANSSEM 700
 FT TRANSSEM 701
 FT TRANSSEM 725
 FT TRANSSEM 745
 FT TRANSSEM 746
 FT TRANSSEM 769
 FT TRANSSEM 770
 FT TRANSSEM 792
 FT TRANSSEM 793
 FT TRANSSEM 805
 FT TRANSSEM 806
 FT TRANSSEM 828
 FT TRANSSEM 829
 FT TRANSSEM 836
 FT TRANSSEM 837
 FT TRANSSEM 862
 FT TRANSSEM 863
 FT TRANSSEM 1079
 FT TRANSSEM 90
 FT TRANSSEM 90
 FT TRANSSEM 130
 FT TRANSSEM 130

FT	CARBOHYD	261	261	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	287	287	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	446	446	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	468	468	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	488	488	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	541	541	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	594	594	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SO	SEQUENCE	1079 AA;	120867 MW;	D7664550361F9736 CRC64;	

Query Match 65.48; Score 34; DB 1; Length 1079;
 Best Local Similarity 62.58; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy	2	ETINCHK 9
Db	354	ETINCHKLO 361

Search completed: July 11, 2003, 11:55:10
 Job time : 2.86071 secs


```

AC 092035;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Megslin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21368006; PubMed=11473647;
RA Nagaoka M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kuokawa K.;
RT "Cloning of rodent megslin revealed its up-regulation in
RT mesangiolipoliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL; AF105329; AAL16769.1; -.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin.1.
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
KM Serpin.
SQ SEQUENCE 380 AA; 42821 MW; D8076CABEE2C2FC6 CRC64;

Query Match 75.0%; Score 39; DB 11; Length 380;
Best Local Similarity 55.6%; Pred. No. 9.1;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SERPINHX 9
Db 188 SDTSCFR 196

RESULT 3
Q9VDU4 PRELIMINARY; PRT; 199 AA.
ID Q9VDU4;
AC Q9VDU4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CG6300 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Laeko P., lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobaric C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzay D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paele J.M.,
RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003727; AAF55696.1; -.
DR HSSP; P08659; ILIC1.
DR FLYbase; FBgn0038730; CG6300.
DR InterPro; IPR001589; Actbind_actln.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding.1.
DR PROSITE; PS00019; ACTININ_1; UNKNOWN_1.
SQ SEQUENCE 199 AA; 22504 MW; DD6AFEC067AFD5DA CRC64;

Query Match 73.1%; Score 38; DB 5; Length 199;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TINCHF 8
Db 3 TINCHF 8

RESULT 4
Q9VDU3 PRELIMINARY; PRT; 208 AA.
ID Q9VDU3;
AC Q9VDU3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CG11659 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Metuliov G., Mishina N.V., Mowbray C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Plitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Sider-Klimos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weisslock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003727; AAF55697.1; -
 DR HSSP: P08659; ILC1.
 DR FLYBASE: FBgn0038731; CG11659.
 DR InterPro: IPR001589; Acbbind_actnln.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00019; ACTININ_1; UNKNOWN_1.
 SQ SEQUENCE 208 AA; 23559 MW; E41495344B417C35 CRC64;

Query Match 73.1%; Score 38; DB 5; Length 208;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TINCH 8
 |||||
 DB 3 TINCH 8

RESULT 5
 092J55 PRELIMINARY; PRT; 224 AA.
 AC 092J55;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein R03183.
 GN R03183 OR SMC03800.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_Taxid=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Bolstead P., Becker A., Boutry M., Cadieu E., Drenon S., Gloux S.,
 RA Godle T., Goffeau A., Kahn D., Kiss E., Lelaune V., Masny D.,
 RA Pohl T., Portetalle D., Puehler A., Purnelle B., Ramseger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT *Sinorhizobium meliloti* strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591793; CAC47762.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 224 AA; 24768 MW; 70D206D4D1AD018 CRC64;

Query Match 71.2%; Score 37; DB 16; Length 224;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ETINCH 7
 |||||
 DB 186 ETINCH 191

RESULT 6
 09JIP9 PRELIMINARY; PRT; 265 AA.
 AC 09JIP9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Lectin-related NK cell receptor LY49L4 (Natural killer cell receptor
 DE LY-49L4).
 GN KLR1A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBA/J;
 RA Makrigianis A.P., Etzler J., Winkler-Pickett R., Ortaldo J.R.,
 RA Anderson S.K.;
 RT "Identification of the LY49L protein: evidence for activating
 RT counterparts to inhibitory receptors.";
 RL J. Leukoc. Biol. 0:0-0(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Silver E.T., Gong D., Hazes B., Kane K.P.;
 RT "Ly-49w: a novel activating receptor of NOD mice with homology to the
 RT inhibitory receptor Ly-49g that recognizes H-2dk and H-2d.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF204268; AAF75837.1; -
 DR EMBL: AF307948; AAG41969.1; -
 DR MGD: MGI:1321091; Klr1a2.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: SM00059; Lectin_c; 1.
 DR SMART: SM00034; CLECT; 1.
 DR PROSITE: PS00041; C-TYPE_LLECTIN_2; 1.
 KW Receptor.
 SQ SEQUENCE 265 AA; 30497 MW; 0F4B927BF1CFEB8 CRC64;

Query Match 71.2%; Score 37; DB 11; Length 265;
 Best Local Similarity 75.0%; Pred. No. 16;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ETINCH 9
 |||||
 DB 77 ETINCHDK 84

RESULT 7
 09EPB2 PRELIMINARY; PRT; 265 AA.
 AC 09EPB2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Natural killer cell receptor LY49W2 (Natural killer cell receptor Ly-
 DE 49W2).
 GN KLR23 OR KLR17 OR LY49W OR LY-49W.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RA "Ly49w: a novel activating receptor of NOD mice with homology to the
 RT inhibitory receptor Ly49g that recognizes H-2dk and H-2d.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOR;

```

RA Silver E.T., Gong D.-E., Hazes B., Kane K.P.;
RT "Ly-49w: a novel activating receptor of NOD mice with homology to the
RT inhibitory receptor Ly-49g that recognizes H-2Dk and H-2Dd.";
RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF283251; AAC10158.1; -.
DR EMBL: AF074463; AAC09059.1; -.
DR MGI: MGI:1930932; K1Ra23.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 265 AA; 30443 MW; 381C38DC9C9A9073 CRC64;

Query Match 71.2%; Score 37; DB 11; Length 265;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ETINCHFK 9
Db 77 ETLNCHDK 84

RESULT 8
O99P19 PRELIMINARY; PRT; 265 AA.
AC O99P19;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Natural killer cell receptor LY49M.
CN K1RA13 OR LY49M.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RA MEDLINE=21103262; PubMed=11160290;
RA Silver E.T., Gong D.-E., Hazes B., Kane K.P.;
RT "Ly-49w, an Activating Receptor of Nonobese Diabetic Mice with Close
RT Homology to the Inhibitory Receptor Ly-49g, Recognizes H-2Dk and H-
RT 2Dd.";
RL J. Immunol. 166:2333-2341(2001).
DR EMBL: AF283252; AAk11559.1; -.
DR MGI: MGI:1321090; K1Ra13.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 265 AA; 30733 MW; F93E1D49EC2670CF CRC64;

Query Match 71.2%; Score 37; DB 11; Length 265;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ETINCHFK 9
Db 77 ETLNCHDK 84

RESULT 9
O9J101 PRELIMINARY; PRT; 268 AA.
AC O9J101;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Lectin-related NK cell receptor LY49L2 (Natural killer cell receptor
DE Ly-49L2).
CN K1RA12.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA/J;
RA Makrigianis A.P., Etzler J., Winkler-Pickett R., Ortaldo J.R.,
RA Anderson S.K.;
RT "Identification of the Ly49L protein: evidence for activating
RT counterparts to inhibitory receptors.";
RL J. Leukoc. Biol. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Silver E.T., Gong D., Hazes B., Kane K.P.;
RT "Ly-49w: a novel activating receptor of NOD mice with homology to the
RT inhibitory receptor Ly-49g that recognizes H-2Dk and H-2Dd.";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF204266; AAF75835.1; -.
DR EMBL: AF307947; AAC41968.1; -.
DR MGI: MGI:1321091; K1Ra12.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 268 AA; 30786 MW; D15013DF7417D4AA CRC64;

Query Match 71.2%; Score 37; DB 11; Length 268;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 ETINCHFK 9
Db 80 ETLNCHDK 87

RESULT 10
O9EP57 PRELIMINARY; PRT; 268 AA.
AC O9EP57;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Natural killer cell receptor LY49W1 (Natural killer cell receptor Ly-
DE 49W1).
CN K1RA23 OR K1RA17 OR LY49W OR LY-49W.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RA Silver E.T., Gong D.-E., Hazes B., Kane K.P.;
RT "Ly-49w: a novel activating receptor of NOD mice with homology to the
RT inhibitory receptor Ly-49g that recognizes H-2Dk and H-2Dd.";
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR;
RA Silver E.T., Gong D.-E., Hazes B., Kane K.P.;
RT "Ly-49w: a novel activating receptor of NOD mice with homology to the
RT inhibitory receptor Ly-49g that recognizes H-2Dk and H-2Dd.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF283250; AAC10157.1; -.
DR EMBL: AF074459; AAC09058.1; -.
DR MGI: MGI:1930932; K1Ra23.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.

```

KW Receptor.
SQ SEQUENCE 268 AA; 30733 MW; E607B9781942AA61 CRC64;
Query Match 71.2%; Score 37; DB 11; Length 268;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 2 ETINCHFK 9
||:||||
Db 80 ETLNCHDK 87
RESULT 11
O9J100 PRELIMINARY; PRT; 278 AA.
AC O9J100; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Lectin-related NK cell receptor LY49LJ3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA/J;
RA Makrigianis A.P., Etzler J., Winkler-Pickett R., Ortaldo J.R.,
RT "Identification of the LY49L protein: evidence for activating
counterparts to inhibitory receptors."
RL J. Leukoc. Biol. 0:0-0(2000).
DR EMBL: AF204267; AAF75836.1; -.
DR MGI: 1321091; KlrA12.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS50041; C_Type_Lectin_2; 1.
KW Receptor.
SQ SEQUENCE 278 AA; 31885 MW; 5F1B09E0FD89EA80 CRC64;
Query Match 71.2%; Score 37; DB 11; Length 278;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 2 ETINCHFK 9
||:||||
Db 77 ETLNCHDK 84
RESULT 12
O9J102 PRELIMINARY; PRT; 281 AA.
AC O9J102; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN Lectin-related NK cell receptor LY49LJ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA/J;
RA Makrigianis A.P., Etzler J., Winkler-Pickett R., Ortaldo J.R.,
RT "Identification of the LY49L protein: evidence for activating
counterparts to inhibitory receptors."
RL J. Leukoc. Biol. 0:0-0(2000).
DR EMBL: AF204265; AAF75834.1; -.
KW Receptor.

DR MGI:1321091; KlrA12.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS50041; C_Type_Lectin_2; 1.
KW Receptor.
SQ SEQUENCE 281 AA; 32175 MW; 2589356CB9F4DE09 CRC64;
Query Match 71.2%; Score 37; DB 11; Length 281;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 2 ETINCHFK 9
||:||||
Db 80 ETLNCHDK 87
RESULT 13
O12180 PRELIMINARY; PRT; 1030 AA.
AC O12180; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Chromosome XV reading frame ORF YOL089C.
OS HAL9 OR YOL089C.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Zumbstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-FY1679;
RX MEDLINE-96021609; PubMed-8533473;
RA Zumbstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains
more than twice as many unknown as known open reading frames."
YL Yeast 11:975-986(1995).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE BINUCLEAR CLUSTER DOMAIN.
DR EMBL: Z74831; CA999101.1; -.
DR EMBL: X83121; CA58190.1; -.
DR HSP; P21228; 2AUC.
DR SGD: S0005449; HAL9.
DR InterPro: IPR001138; Fungi_trn.
DR Pfam: PF00172; zn_c1us; 1.
DR SMART: SM00066; GAL4; 1.
DR PROSITE: PS00463; ZN2_CV6_FUNGAL_1; 1.
DR PROSITE: PS50048; ZN2_CV6_FUNGAL_2; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription regulation;
KW Zinc.
SQ SEQUENCE 1030 AA; 117925 MW; 8C8BDE8406D2105 CRC64;
Query Match 71.2%; Score 37; DB 3; Length 1030;
Best Local Similarity 55.0%; Pred. No. 54;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 SETINCHFK 9
|:||||:
Db 689 SRVLNCHR 697
RESULT 14
O9V9N7 PRELIMINARY; PRT; 1354 AA.
ID O9V9N7

```

AC 09Y9N7;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)
DE CG6448 protein.
GN CG6448.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adamatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Beeson P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu L.B., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieri S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.N., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003782; AAF57249.1; -
DR FLYBASE: FBgn0032976; CG6448.
SQ SEQUENCE 1354 AA; 150276 MW; 770F5904A634EFD6 CRC64;

```

```

Query Match          69.2%; Score 36; DB 5; Length 1354;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 ETINCH 7
DB 1172 ETVNCN 1177

```

```

RESULT 15
ID 08S6H9 PRELIMINARY; PRT: 1739 AA.
AC 08S6H9;
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)

```

```

DE Putative retroelement.
GN OSJNB0023M1.15.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euharoidae; Oryzae; Oryza.
RX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTIPONBARE;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nascimento L., Baker J., Santos L., Zutavern T., Miller B.,
RA Cunniss D.M., Katzenberger F., Muller S., Bell M., Ballija V., Shah R.,
RA King L., Yang C., Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNB0023M1, from chromosome 10, complete sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC092749; AAM08361.1; -
SQ SEQUENCE 1739 AA; 196006 MW; 47F7707D33152286 CRC64;

```

```

Query Match          69.2%; Score 36; DB 10; Length 1739;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
OY 2 ETINCH 9
DB 1035 QSVNCHQ 1042

```

```

Search completed: July 11, 2003, 11:58:48
Job time : 5.92931 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 4.01663 Seconds
(without alignments)
232.223 Million cell updates/sec

Title: US-10-091-442-2
Perfect score: 40
Sequence: 1 WQSAFTK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	380	15	AAK48379
2	40	100.0	380	20	AAV08254
3	40	100.0	380	21	AAK24142
4	40	100.0	380	22	AAK83075
5	36	90.0	380	22	ABG02777
6	36	90.0	323	22	ABG11497
7	36	90.0	368	20	AAV06236
8	36	90.0	368	21	AAK24151
9	36	90.0	368	22	AAK83077
10	36	90.0	380	20	AAV08255

11	36	90.0	380	21	AAK24150
12	36	90.0	380	22	AAK64286
13	36	90.0	380	22	AAK83076
14	32	80.0	219	22	ABG29897
15	32	80.0	219	22	ABG18342
16	32	80.0	931	21	AAV92833
17	32	80.0	938	23	ABK90528
18	32	80.0	940	20	AAV35082
19	32	80.0	944	21	AAK43351
20	32	80.0	959	22	AAK64950
21	32	80.0	1072	22	ABK62164
22	32	77.5	46	21	AAK28753
23	31	77.5	119	22	AAO06492
24	31	77.5	121	21	AAK58546
25	31	77.5	125	22	AAK91461
26	31	77.5	168	22	ABK10348
27	31	77.5	187	21	AAK37074
28	31	77.5	187	21	AAK52578
29	31	77.5	190	21	AAK52577
30	31	77.5	191	21	AAK37073
31	31	77.5	193	20	AAK38921
32	31	77.5	194	22	AAU18012
33	31	77.5	202	21	AAK06264
34	31	77.5	264	20	AAV35068
35	31	77.5	356	20	AAV38925
36	31	77.5	357	20	AAV38922
37	31	77.5	357	20	AAV38923
38	31	77.5	364	20	AAV38924
39	31	77.5	514	22	AAK96264
40	31	77.5	666	23	ABP25064
41	31	77.5	733	20	AAV35373
42	31	77.5	813	21	AAV92835
43	31	77.5	978	21	AAV92834
44	31	77.5	978	22	AAV38886
45	31	77.5	978	23	ABK90543

ALIGNMENTS

RESULT 1	AAK48379	standard; Protein; 380 AA.
ID	AAK48379	
AC	AAK48379;	
XX		
DT	16-AUG-1994 (first entry)	
XX		
DE	Human megakaryocyte differentiation factor.	
XX		
KW	Human megakaryocyte differentiation factor; MDF; thrombopoietin;	
KW	hematopoietic stimulating factor; thrombocytopoietin; platelet;	
KW	bone marrow transplantation; cancer chemotherapy.	
XX		
OS	Homo sapiens.	
XX		
PN	EP583884-A.	
XX		
PD	23-FEB-1994.	
XX		
PF	19-JUL-1993; 93EP-0305654.	
XX		
PR	17-JUL-1992; 92JP-0212305.	
XX		
PA	04-MAR-1993; 93JP-0067339.	
XX		
PA	(SUNR) SUNTORY LTD.	
XX	(TSUJ) TSUJIMOTO M.	
XX		
PI	Ishida N, Iwasa F, Kurihara T, Mura K, Nakazato H;	
XX	TSUJIMOTO M, TSURUOKA N, Yamaguchi N, Yamachi K;	
DR	WPI: 1994-058782/08.	
DR	N-PSDB: AAQ56670.	

Rat megilin protein
Rat megilin protein
Rat megilin protein
Novel human diseno
Novel human diseno
C. pneumoniae CPN1
Chlamydia pneumonia
Human ORF3115
Human ORF3115
Human ORF3115
Drosophila melanog
Protein fragment e
Human polyepitide
Arabidopsis thalia
C. glutamicum prote
Human cDNA SEQ ID
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Neisseria meningit
Human immunoglobul
Arabidopsis thalia
Chlamydia pneumoni
Neisseria gonorrhoe
Neisseria meningit
Neisseria meningit
Neisseria gonorrhoe
Putative P. abyssal
Streptococcus poly
Amino acid sequenc
C. pneumoniae CPN1
C. pneumoniae CT87
Chlamydia pneumoni

```

XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
XX decrease in platelets
PS Claim 7; Page 30-32; 47pp; English.
XX
CC Human MDF can be isolated from a culture of human epidermoid
CC carcinoma A431 cells in protein-free medium. The MDF stimulates
CC differentiation of megakaryocytes from myeloid cells in the presence
CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC for treatment of diseases involving a decrease in platelet number
CC (esp. thrombocytopaenia) such as occurs in bone marrow
CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kd by
CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 40; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
DB 181 WQSAFTK 187

RESULT 2
AA08254
ID AA08254 standard; Protein: 380 AA.
XX
AC AAY08254;
XX
DT 14-JUL-1999 (first entry)
XX
DE Human megsin protein.
XX
KW Megsin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
XX human; rat; murine.
XX
OS Homo sapiens.
XX
PN W09915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
XX (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 1999-276983/23.
DR N-PSDB; AAX56712.
XX
PT Megsin protein expressed specifically in mesangial cells
XX
PS Claim 1; Page 62-64; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megsin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as Iga nephropathy.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 40; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 WQSAFTK 7
DB 181 WQSAFTK 187

RESULT 3
AAB24142
ID AAB24142 standard; Protein: 380 AA.
XX
AC AAB24142;
XX
DT 30-JAN-2001 (first entry)
XX
DE Human megsin protein sequence SEQ ID NO:2.
XX
KW Megsin; mesangium-predominant gene; seipin regulated; nephropathy;
KW Iga; immunoglobulin A; detection; renal function; renal disorder;
XX diagnosis; biological sample; blood; urine.
XX
OS Homo sapiens.
XX
PN W0200057189-A1.
XX
PD 28-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-JP01646.
XX
PR 19-MAR-1999; 99JP-0075305.
XX 28-OCT-1999; 99JP-0306623.
XX
PA (KURO/) KUROKAWA K.
XX (FUSO ) FUSO PHARM IND LTD.
XX (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 2000-611642/58.
DR N-PSDB; AAA99294.
XX
PT Evaluating renal function comprises assaying megsin protein in
XX biological sample
XX
PS Example 2; Page 66-69; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megsin protein in biological sample. Also
CC described are: (1) use of a anti-megsin protein antibody for diagnosing
CC renal function; and (2) a kit for detecting megsin protein comprising:
CC (a) anti-megsin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying megsin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence represents the human megsin
CC protein, which is given in the exemplification of the present invention.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 40; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
DB 181 WQSAFTK 187

RESULT 4
AAB83075
ID AAB83075 standard; Protein: 380 AA.
XX
AC AAB83075;

```


XX 10-JUL-2001 (first entry)
 XX
 DE Human mesgln protein.
 XX
 KW Human: mesgln; mesangial cell proliferative nephritis; nephrotoxic;
 KW transgenic mouse; glomerular disease; animal model; drug screening.
 XX
 OS Homo sapiens.
 XX
 PM WO200124628-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 06-OCT-2000; 2000WO-JP06988.
 XX
 PR 06-OCT-1999; 99JP-0285736.
 XX
 PA (KUROO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2001-300136/31.
 DR N-PSDB: AAF82438.
 XX
 PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments
 XX
 PS Example 4; Page 44-46; 62pp; Japanese.
 XX
 CC The present sequence is human mesgln. The human mesgln coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analyzing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.
 XX
 SO Sequence 380 AA;
 XX
 Query Match 100.0%; Score 40; DB 22; Length 380;
 Best Local Similarity 100.0%; Pred. NO. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 WQSAFTK 7
 |||||
 DB 181 WQSAFTK 187

RESULT 5
 ABG02777
 ID ABC02777 standard; Protein: 306 AA.
 XX
 AC ABG02777;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #2768.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PM WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX

PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS66964.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 33136; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 306 AA;
 XX
 Query Match 90.0%; Score 36; DB 22; Length 306;
 Best Local Similarity 85.7%; Pred. NO. 54;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 WQSAFTK 7
 |||||
 DB 108 WQSVFTK 114

RESULT 6
 ABG11497
 ID ABG11497 standard; Protein: 323 AA.
 XX
 AC ABG11497;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #11488.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PM WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX

XX (HXSE-) HXSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR N-PSDB: AAS75684.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS
 PS Claim 20: SEQ ID No 41856; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridization probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 323 AA;
 OY 1 WQSAFTK 7
 Db 108 WQSVFTK 114
 Query Match 90.0%; Score 36; DB 22; Length 323;
 Best Local Similarity 85.7%; Pred. No. 57;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 RESULT 7
 AAY08256
 ID AAY08256 standard; Protein: 368 AA.
 XX
 AC AAY08256;
 XX
 DT 14-JUL-1999 (first entry)
 XX
 DE Mouse megalin protein.
 XX
 KM Megalin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;
 KM human; rat; murine.
 XX
 OS Mus musculus.
 XX
 FT Key Location/Qualifiers
 FT Protein 1..368
 FT /note= "partial sequence"
 XX
 PN W09915652-A1.
 PD 01-APR-1999.
 XX
 PF 22-SEP-1998; 98WO-JP04269.
 XX

PR 22-SEP-1997; 97JP-0275302.
 XX
 PA (KURO/) KUROKAWA K.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 1999-276983/23.
 DR N-PSDB: AAX56714.
 XX
 PT Megalin protein expressed specifically in mesangial cells
 PT
 PS Claim 1: Page 76-79; 100pp; Japanese.
 CC This invention describes the isolation of novel megalin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as IGA nephropathy.
 CC
 SQ Sequence 368 AA;
 OY 1 WQSAFTK 7
 Db 169 WKSFTK 175
 Query Match 90.0%; Score 36; DB 20; Length 368;
 Best Local Similarity 85.7%; Pred. No. 65;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 RESULT 8
 AAB24151
 ID AAB24151 standard; Protein: 368 AA.
 XX
 AC AAB24151;
 XX
 DT 30-JAN-2001 (first entry)
 XX
 DE Mouse megalin protein sequence SEQ ID NO:21.
 XX
 KM Megalin; mesangium-predominant gene; serpin regulated; nephropathy;
 KM IGA; immunoglobulin A; detection; renal function; renal disorder;
 KM diagnosis; biological sample; blood; urine.
 XX
 OS Mus musculus.
 XX
 PN W0200057189-A1.
 PD 28-SEP-2000.
 XX
 PE 17-MAR-2000; 2000WO-JP01646.
 XX
 PR 19-MAR-1999; 99JP-0075305.
 PR 28-OCT-1999; 99JP-0306623.
 XX
 PA (KURO/) KUROKAWA K.
 PA (FUSO) FUSO PHARM IND LTD.
 PA (MIYA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2000-611642/58.
 DR N-PSDB: AAC55239.
 XX
 PT Evaluating renal function comprises assaying megalin protein in
 PT biological sample
 XX
 PS Disclosure: Page 89-91; 93pp; Japanese.
 CC The present invention describes a method for evaluating renal function.
 CC The method comprises assaying megalin protein in biological sample. Also
 CC described are: (1) use of a anti-megalin protein antibody for diagnosing

CC renal function; and (2) a kit for detecting mesgin protein comprising:
 CC (a) anti-mesgin protein antibody attached to solid magnetic particles;
 CC (b) direct or indirect fixing for the antibody to the particles; and
 CC (c) a magnet. The process is useful for evaluating renal function and
 CC diagnosing renal disorders by assaying mesgin protein in biological
 CC samples (preferably urine or blood). The process is reproducible and
 CC gives accurate results. The present sequence represents the mouse mesgin
 CC protein, which is given in the exemplification of the present invention.
 XX

SO Sequence 368 AA:

Query Match 90.0%; Score 36; DB 21; Length 368;
 Best Local Similarity 85.7%; Pred. No. 65;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
 1:|||||
 Db 169 WKSFTK 175

RESULT 9

ID AAB83077 standard; Protein: 368 AA.

AC AAB83077;

DT 10-JUL-2001 (first entry)

DE Murine mesgin protein.

KW Mouse: mesangial cell proliferative nephritis; nephrotropic;
 transgenic mouse; glomerular disease; animal model; drug screening.

OS Mus musculus.

PN WO200124628-A1.

PD 12-APR-2001.

PE 06-OCT-2000; 2000WO-JP06988.

PR 06-OCT-1999; 99JP-0285736.

PA (KUROO/) KUROKAWA K.

PA (MIYA/) MIYATA T.

PI Miyata T;

DR WPI: 2001-300136/31.

DR N-PSDB: AAF82440.

PT Mouse model for mesangial cell proliferative nephritis for development
 and screening of new treatments

PS Disclosure: Page 52-53; 62pp; Japanese.

CC The present sequence is murine mesgin protein. The human mesgin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful
 CC for analyzing the pathology of chronic glomerular diseases and for
 CC screening compositions for prevention and treatment of the diseases.
 CC Highly uniform models can be made easily and in large numbers using
 CC this method.

SO Sequence 368 AA:

Query Match 90.0%; Score 36; DB 22; Length 368;
 Best Local Similarity 85.7%; Pred. No. 65;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7

Db 1:|||||
 169 WKSFTK 175

RESULT 10

ID AAY08255 standard; Protein: 380 AA.

AC AAY08255;

DT 14-JUL-1999 (first entry)

DE Rat mesgin protein.

KW Mesgin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;
 human; rat; murine.

OS Rattus rattus.

PN WO9915652-A1.

PD 01-APR-1999.

PE 22-SEP-1998; 98WO-JP04269.

PR 22-SEP-1997; 97JP-0275302.

PA (KUROO/) KUROKAWA K.

PA (MIYA/) MIYATA T.

PI Miyata T;

DR WPI: 1999-276983/23.

DR N-PSDB: AAX56712.

PT Mesgin protein expressed specifically in mesangial cells

PS Claim 1; Page 69-72; 100pp; Japanese.

CC This invention describes the isolation of novel mesgin nucleic acid and
 CC proteins from human, rat and mouse tissue. This protein is expressed
 CC specifically in mesangial cells. The products of the invention are
 CC useful for the treatment and diagnosis of diseases involving mesangial
 CC cells, such as Iga nephropathy.

SO Sequence 380 AA:

Query Match 90.0%; Score 36; DB 20; Length 380;
 Best Local Similarity 85.7%; Pred. No. 67;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
 1:|||||
 Db 181 WKSFTK 187

RESULT 11

ID AAB24150 standard; Protein: 380 AA.

AC AAB24150;

DT 30-JAN-2001 (first entry)

DE Rat mesgin protein sequence SEQ ID NO:19.

KW Mesgin; mesangium-predominant gene; serpin regulated; nephropathy;
 Iga; immunoglobulin A; detection; renal function; renal disorder;
 diagnosis; biological sample; blood; urine.
 OS Rattus norvegicus.

FH key Location/Qualifiers

FT	Misc-difference	51	/note= "unspecified"
FT	Misc-difference	94	/note= "unspecified"
FT	Misc-difference	94	/note= "unspecified"
XX	WO200057189-A1.		
XX	28-SEP-2000.		
XX	17-MAR-2000; 2000WO-JP01646.		
XX	19-MAR-1999; 99JP-0075305.		
PR	28-OCT-1999; 99JP-0306623.		
XX			
PA	(KURO/ KUROKAWA K.		
PA	(FUSO) FUSO PHARM IND LTD.		
XX	(MIYA/) MIYATA T.		
PI	Miyata T;		
XX			
DR	WPI: 2000-611642/58.		
DR	N-PSDB; AAC55238.		
XX			
PT	Evaluating renal function comprises assaying megin protein in		
PT	biological sample		
XX			
PS	Example 2; Page 81-84; 93pp; Japanese.		
XX			
CC	The present invention describes a method for evaluating renal function.		
CC	The method comprises assaying megin protein in biological sample. Also		
CC	described are: (1) use of a anti megin protein antibody for diagnosing		
CC	renal function; and (2) a kit for detecting megin protein comprising:		
CC	(a) anti megin protein antibody attached to solid magnetic particles;		
CC	(b) direct or indirect fixing for the antibody to the particles; and		
CC	(c) a magnet. The process is useful for evaluating renal function and		
CC	diagnosing renal disorders by assaying megin protein in biological		
CC	samples (preferably urine or blood). The process is reproducible and		
CC	gives accurate results. The present sequence represents the rat megin		
CC	protein, which is given in the exemplification of the present invention.		
XX			
XX			
XX	Sequence 380 AA;		
XX			
XX			
XX	Query Match	90.0%;	Score 36; DB 21; Length 380;
XX	Best Local Similarity	85.7%;	Pred. No. 67;
XX	Matches 6; Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
OY	1 WOSAFTK 7		
DB	181 WKSAPTK 187		
XX			
XX	RESULT 12		
XX	AAG64286		
XX	AAG64286 standard; Protein; 380 AA.		
XX	AAG64286;		
XX	21-SEP-2001 (first entry)		
XX			
XX	Rat megin protein.		
XX			
XX	Rat; megin; renal mesangial cell; mesangium proliferative nephritis.		
XX	Rattus norvegicus.		
XX	WO200148019-A1.		
XX			
XX	05-JUL-2001.		
XX			
XX	26-DEC-2000; 2000WO-JP09251.		
XX			
XX	28-DEC-1999; 99JP-0373677.		
XX			

PA (KUROO/) KUROKAWA K.
 PA (MIRA/) MIYATA T.
 XX Miyata T;
 PI
 XX WPI: 2001-425651/45.
 DR N-PSDB: AAH48181.
 XX
 PT New antibody recognizing a partial sequence of rat megin protein for
 diagnosis of mesangium proliferative nephritis -
 XX
 PS Disclosure: Page 54-56; 63pp; Japanese.
 XX
 CC The present invention relates to a novel antibody which recognises a
 CC peptide consisting of residues 341-354 of rat megin protein. The present
 CC sequence is the protein sequence for rat megin, which was used in the
 CC present invention. Megsin is highly expressed in renal mesangial cells
 CC and its level is elevated in mesangium proliferative nephritis. Assay of
 CC the serum or urine level using the antibody is therefore indicative of
 CC this type of disorder.
 XX
 SQ Sequence 380 AA;
 QY
 Db 1 WSAFTK 7
 1:|||||
 181 WKSFTK 187
 RESULT 13
 AAB83076
 ID AAB83076 standard; Protein; 380 AA.
 XX
 AC AAB83076;
 XX
 DT 10-JUL-2001 (first entry)
 XX
 DE Rat megin protein.
 XX
 KW Rat; megin; mesangial cell proliferative nephritis; nephrotropic;
 KW Transgenic mouse; glomerular disease; animal model; drug screening.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200124628-A1.
 PD 12-APR-2001.
 XX
 PF 06-OCT-2000; 2000WO-JP06988.
 XX
 PR 06-OCT-1999; 99JP-0285736.
 XX
 PA (KUROO/) KUROKAWA K.
 PA (MIRA/) MIYATA T.
 XX
 PI Miyata T;
 XX
 DR WPI: 2001-300136/31.
 DR N-PSDB: AAF82439.
 XX
 PT Mouse model for mesangial cell proliferative nephritis for development
 PT and screening of new treatments -
 XX
 PS Disclosure: Page 48-50; 62pp; Japanese.
 CC The present sequence is rat megin. The human megin coding
 CC sequence may be introduced into a mouse to produce an animal model of
 CC mesangial cell proliferative nephritis. The symptoms include
 CC enlargement of the mesangial base region, sedimentation of an immune
 CC complex and an increase in mesangial cells. The animal model is useful

CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.

XX Sequence 380 AA;

Query Match 90.0%; Score 36; DB 22; Length 380;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
1:|||||
Db 181 WKSATFK 187

RESULT 14
ABG29897
ID ABG29897 standard; Protein; 219 AA.

XX AC ABG29897;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #29888.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS94084.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 60256; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 219 AA;

Query Match 80.0%; Score 32; DB 22; Length 219;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
1:|||||
Db 196 WESEFTK 202

RESULT 15
ABG18342
ID ABG18342 standard; Protein; 331 AA.

XX AC ABG18342;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #18331.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB; AAS82529.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 48701; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SO Sequence 331 AA;

Query Match 80.0%; Score 32; DB 22; Length 331;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
- 1:1 111
Db 308 WESEFTK 314

Search completed: July 11, 2003, 11:54:18
Job time : 5.01663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 1.25156 seconds

(without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-2

Perfect score: 40

Sequence: 1 WQSAFTK 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/Dackfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	7	2	US-08-611-977-2
2	40	100.0	380	2	US-08-472-659-34
3	40	100.0	380	2	US-08-474-661-34
4	40	100.0	380	2	US-08-611-977-34
5	30	75.0	385	3	US-09-053-866-2
6	30	75.0	385	4	US-09-479-130-2
7	29	72.5	381	4	US-09-216-295-22
8	28	70.0	146	1	US-08-543-238-11
9	28	70.0	146	1	US-08-420-526-11
10	28	70.0	177	4	US-09-134-001C-4482
11	28	70.0	189	2	US-08-483-695-46
12	28	70.0	189	2	US-07-965-285-46
13	28	70.0	189	2	US-08-487-231-46
14	28	70.0	189	4	US-09-201-912-46
15	28	70.0	269	4	US-09-045-973-3
16	28	70.0	269	2	US-07-857-224B-48
17	28	70.0	412	4	US-09-134-001C-4885
18	28	70.0	452	1	US-08-275-488A-2
19	28	70.0	452	1	US-08-275-490-2
20	28	70.0	452	1	US-08-446-380-2
21	28	70.0	452	1	US-08-446-374-2
22	28	70.0	452	1	US-08-446-382-2
23	28	70.0	452	1	US-08-445-801-2
24	28	70.0	452	1	US-08-275-487-2
25	28	70.0	452	5	PCR-US95-08919-2
26	28	70.0	453	1	US-08-275-488A-12
27	28	70.0	453	1	US-08-275-490-12

28	28	70.0	453	1	US-08-446-380-12	Sequence 12, Appl
29	28	70.0	453	1	US-08-446-374-12	Sequence 12, Appl
30	28	70.0	453	1	US-08-446-382-12	Sequence 12, Appl
31	28	70.0	453	1	US-08-445-801-12	Sequence 12, Appl
32	28	70.0	453	1	US-08-275-487-12	Sequence 12, Appl
33	28	70.0	453	5	PCR-US95-08919-12	Sequence 12, Appl
34	28	70.0	494	1	US-08-275-490-4	Sequence 4, Appl1
35	28	70.0	494	1	US-08-446-380-4	Sequence 4, Appl1
36	28	70.0	494	1	US-08-446-374-4	Sequence 4, Appl1
37	28	70.0	494	1	US-08-446-382-4	Sequence 4, Appl1
38	28	70.0	494	1	US-08-445-801-4	Sequence 4, Appl1
39	28	70.0	494	1	US-08-275-487-4	Sequence 4, Appl1
40	28	70.0	494	5	PCR-US95-08919-4	Sequence 4, Appl1
41	28	70.0	609	1	US-08-324-977-40	Sequence 40, Appl
42	28	70.0	609	2	US-08-384-616-40	Sequence 40, Appl
43	28	70.0	609	2	US-08-904-686A-40	Sequence 40, Appl
44	28	70.0	609	4	US-09-315-850-40	Sequence 40, Appl
45	28	70.0	609	4	US-09-315-850-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-611-977-2
; Sequence 2, Application US/08611977
; Patent No. 5972866
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. 5972886uo
APPLICANT: NAKAZOTO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-2

Query Match 100.0%; Score 40; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
1111111
DB 1 WOSAFTK 7

RESULT 2
US-08-472-659-34

Sequence 34, Application US/08472659

Patent No. 5831030

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5831030uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5831030unhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMACHII, Kozo

APPLICANT: YAMAGUCHI, No. 5831030om1

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472, 659

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300

REFERENCE/DOCKET NUMBER: 001560-248

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-472-659-34

Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
1111111
DB 181 WOSAFTK 187

RESULT 3
US-08-474-661-34

Sequence 34, Application US/08474661

Patent No. 5874253

GENERAL INFORMATION:

APPLICANT: TSUJIMOTO, Masafumi

APPLICANT: IWASA, Fuyuki

APPLICANT: TSURUOKA, No. 5874253uo

APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju

APPLICANT: ISHIDA, No. 5874253unhiro

APPLICANT: KURIHARA, Tatsuya

APPLICANT: YAMACHII, Kozo

APPLICANT: YAMAGUCHI, No. 5874253om1

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: George Mason Bldg., Washington & Prince Sts.

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,661

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-212305

FILING DATE: 17-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-067339

FILING DATE: 04-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: REA, TERESA STANER

REGISTRATION NUMBER: 30,427

REFERENCE/DOCKET NUMBER: 001560-204

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6620

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 380 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-474-661-34

Query Match 100.0%; Score 40; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
1111111
DB 181 WOSAFTK 187

RESULT 4

US-08-611-977-34
; Sequence 34, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUTSUMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSURODOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhitro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886om1
; TITLE OF INVENTION: MECAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-611-977-34

Query Match 100.0%; Score 40; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WOSAFK 7
1111111
DB 181 WOSAFK 187

RESULT 5
US-09-053-866-2
; Sequence 2, Application US/09053866
; Patent No. 6111075
; GENERAL INFORMATION:
; APPLICANT: Xu, Wenfeng
; APPLICANT: Presnell, Scott R.
; APPLICANT: Yee, David P.
; APPLICANT: Foster, Donald C.

TITLE OF INVENTION: PROTEASE-ACTIVATED RECEPTOR
; TITLE OF INVENTION: PAR4 (ZCHEMR2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leith, Debra K
; REGISTRATION NUMBER: 32,619
; REFERENCE/DOCKET NUMBER: 98-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-053-866-2

Query Match 75.0%; Score 30; DB 3; Length 385;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WOSAFK 6
111111
DB 241 WOSAFK 246

RESULT 6
US-09-479-130-2
; Sequence 2, Application US/09479130
; Patent No. 6436400
; GENERAL INFORMATION:
; APPLICANT: Xu, Wenfeng
; APPLICANT: Presnell, Scott R.
; APPLICANT: Yee, David P.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: PROTEASE-ACTIVATED RECEPTOR
; TITLE OF INVENTION: PAR4 (ZCHEMR2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/479,130
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Leitch, Debra K
REGISTRATION NUMBER: 32,619
REFERENCE/DOCKET NUMBER: 98-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6674
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-479-130-2

Query Match 75.0%; Score 30; DB 4; Length 385;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFT 6
11111
DB 241 WOLAF 246

RESULT 7
US-09-216-295-22
Sequence 22, Application US/09216295
Patent No. 6268328
GENERAL INFORMATION:
APPLICANT: Mitchinson, Colin
APPLICANT: Mendt, Dan J.
TITLE OF INVENTION: No. 6268328el Variant Egit-Like Cellulase Compositions
FILE REFERENCE: GC555
CURRENT APPLICATION NUMBER: US/09/216,295
CURRENT FILING DATE: 1998-12-18
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22
LENGTH: 381
TYPE: PRT
ORGANISM: Streptomyces lividans CelB
US-09-216-295-22

Query Match 72.5%; Score 29; DB 4; Length 381;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFT 6
11111
DB 309 WOLAF 314

RESULT 8
US-08-543-238-11
Sequence 11, Application US/08543238
Patent No. 5607919
GENERAL INFORMATION:
APPLICANT: Bojsen, Kirsten
APPLICANT: Kragh, Karsten M.
APPLICANT: Mikkelsen, Jorn D.
APPLICANT: Nielsen, Klaus K.
TITLE OF INVENTION: Anti-Microbial Proteins
NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/543,238
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Marcus-Wyner, Lynn
REGISTRATION NUMBER: 34,869
REFERENCE/DOCKET NUMBER: 137-1078/MA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3588
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-543-238-11

Query Match 70.0%; Score 28; DB 1; Length 146;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 WOSAFT 7
11111
DB 116 WDFVFTK 122

RESULT 9
US-08-420-526-11
Sequence 11, Application US/08420526
Patent No. 5608151
GENERAL INFORMATION:
APPLICANT: Bojsen, Kirsten
APPLICANT: Kragh, Karsten M.
APPLICANT: Mikkelsen, Jorn D.
APPLICANT: Nielsen, Klaus K.
TITLE OF INVENTION: Anti-Microbial Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sandoz Agro, Inc.
STREET: 975 California Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/420,526
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Marcus-Wyner, Lynn
REGISTRATION NUMBER: 34,869
REFERENCE/DOCKET NUMBER: 137-1078/MA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/354-3588

TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-420-526-11

Query Match
Best Local Similarity 70.0%; Score 28; DB 1; Length 146;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MOSAFT 7
1 1 1 1 1
DB 116 MDVFTK 122

RESULT 10
US-09-134-001C-4482
Sequence 4482, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4482
LENGTH: 177
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4482

Query Match
Best Local Similarity 70.0%; Score 28; DB 4; Length 177;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOSAFT 6
1 1 1 1 1
DB 88 WMSAFT 93

RESULT 11
US-08-483-695-46
Sequence 46, Application US/08483695
Patent No. 5866139
GENERAL INFORMATION:
APPLICANT: Brechot, Christian
APPLICANT: Kremsdorf, Dina
TITLE OF INVENTION: Porchon, Colette
TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,695
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/965,285
FILING DATE: 18-MAR-1993
APPLICATION NUMBER: FR 91 06 882
FILING DATE: 06-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400

Query Match
Best Local Similarity 70.0%; Score 28; DB 2; Length 189;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOSAFT 6
1 1 1 1 1
DB 104 MESVFT 109

RESULT 12
US-07-965-285-46
Sequence 46, Application US/07965285
Patent No. 5879904
GENERAL INFORMATION:
APPLICANT: Brechot, Christian
APPLICANT: Kremsdorf, Dina
APPLICANT: Porchon, Colette
TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/965,285
FILING DATE: 18-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 06 882
FILING DATE: 06-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-965-285-46

Query Match 70.0%; Score 28; DB 2; Length 189;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WSAFT 6
DB 104 WESVFT 109

RESULT 13
US-08-487-231-46

Sequence 46, Application US/08487231
Patent No. 5919454

GENERAL INFORMATION:
APPLICANT: Brechot, Christian

APPLICANT: Kremsdorf, Dina

TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
Hepatitis C Virus Isolate, Diagnostic and Therapeutic

TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.

CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,231
FILING DATE: 07-JUNE-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/965,285
FILING DATE: 18-MAR-1993

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 91 06 882
FILING DATE: 06-JUN-1991

ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-02000

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-487-231-46

Query Match 70.0%; Score 28; DB 2; Length 189;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WSAFT 6

DB 104 WESVFT 109

RESULT 14
US-09-201-912-46

Sequence 46, Application US/09201912
Patent No. 6210962

GENERAL INFORMATION:
APPLICANT: Brechot, Christian

APPLICANT: Kremsdorf, Dina

TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
Hepatitis C Virus Isolate, Diagnostic and Therapeutic

TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic

NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.

CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/201,912
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/965,285
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05286-0001-00000

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-09-201-912-46

Query Match 70.0%; Score 28; DB 4; Length 189;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WSAFT 6
DB 104 WESVFT 109

RESULT 15
US-09-045-973-3

Sequence 3, Application US/09045973
Patent No. 615767

GENERAL INFORMATION:
APPLICANT: Lal, Preeti

APPLICANT: Yue, Henry

APPLICANT: Corley, Neil C.

APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT18
CLONE: 2534680
US-09-045-973-3

Query Match 70.0%; Score 28; DB 4; Length 261;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WOSAF 6
1:1111
DB 118 WRAFT 123

Search completed: July 11, 2003, 12:02:05
Job time : 2.25156 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 2.05198 Seconds
(without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-2

Perfect score: 40

Sequence: 1 MOSAFWK 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCNT_NEM_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/PCRTS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	7	9	US-10-091-442-2
2	40	100.0	7	9	US-10-140-719-2
3	40	100.0	380	9	US-10-091-442-34
4	40	100.0	380	10	US-09-140-719-34
5	31	77.5	125	9	US-09-738-626-5215
6	31	77.5	168	10	US-09-764-853-656
7	31	77.5	194	9	US-10-091-438-157
8	31	77.5	198	12	US-10-097-693-65
9	30	75.0	195	9	US-10-097-065-172
10	30	75.0	385	9	US-10-225-567A-516
11	30	75.0	687	9	US-09-738-626-6127
12	29	72.5	145	9	US-10-204-887-158
13	29	72.5	165	9	US-09-738-626-6151
14	29	72.5	225	9	US-09-738-626-4605
15	29	72.5	307	9	US-09-854-133-397
16	29	72.5	307	9	US-10-144-649A-397
17	29	72.5	307	10	US-09-738-973-397
18	29	72.5	576	9	US-09-364-847-37
19	29	72.5	693	9	US-10-290-078-9

20	29	72.5	712	9	US-09-364-847-49	Sequence 49, Appl
21	29	72.5	712	9	US-09-364-847-51	Sequence 51, Appl
22	29	72.5	1203	9	US-10-097-340-43	Sequence 43, Appl
23	28	70.0	95	9	US-09-899-046-30	Sequence 30, Appl
24	28	70.0	95	9	US-09-878-281-30	Sequence 30, Appl
25	28	70.0	131	10	US-09-814-122-70	Sequence 70, Appl
26	28	70.0	209	9	US-09-899-046-223	Sequence 223, App
27	28	70.0	209	9	US-09-878-281-223	Sequence 223, App
28	28	70.0	224	9	US-10-156-761-7796	Sequence 7796, Ap
29	28	70.0	243	9	US-09-738-626-6003	Sequence 6003, Ap
30	28	70.0	382	9	US-10-078-770-66	Sequence 66, Appl
31	28	70.0	429	10	US-09-815-242-14022	Sequence 14022, A
32	28	70.0	461	9	US-09-925-299-957	Sequence 957, App
33	28	70.0	461	9	US-10-106-658-5254	Sequence 5254, Ap
34	28	70.0	461	10	US-09-925-299-957	Sequence 957, App
35	28	70.0	503	9	US-10-142-231-63	Sequence 63, Appl
36	28	70.0	631	9	US-10-214-932-78	Sequence 78, Appl
37	28	70.0	768	9	US-10-156-761-13830	Sequence 13830, A
38	28	70.0	1692	9	US-09-919-901-4	Sequence 4, Appl1
39	28	70.0	1692	9	US-09-919-901-11	Sequence 11, Appl
40	28	70.0	1692	9	US-09-919-901-18	Sequence 18, Appl
41	28	70.0	1985	9	US-10-258-275-42	Sequence 42, Appl
42	28	70.0	2201	9	US-10-085-476-2	Sequence 2, Appl1
43	28	70.0	2201	12	US-10-029-907-3	Sequence 3, Appl1
44	28	70.0	2307	9	US-09-919-901-2	Sequence 2, Appl1
45	28	70.0	2307	9	US-09-919-901-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1

US-10-091-442-2

Sequence 2, Application US/10091442

Patent No. US20020164711A1

GENERAL INFORMATION:

APPLICANT: TSUTSUMOTO, Masafumi

IMASA, Fuyuki

TSURUOKA, No. US20020164711A1uo

NAKAZATO, Hirosaki

MURA, Kenju

ISHIDA, No. US20020164711A1uhfro

KURIHARA, Tatsuya

YAMAGUCHI, Kozo

YAMAGUCHI, No. US20020164711A1omf

TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Burns, Doane, Swecker & Mathis

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/091,442

FILING DATE: 07-Mar-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/140,719

FILING DATE: 08-AUG-1998

APPLICATION NUMBER: US 08/474,661

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/091,028

FILING DATE: 14-JUL-1993

APPLICATION NUMBER: JP 4-213305

FILING DATE: 17-JUL-1992

APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-091-442-2

Query Match 100.0%; Score 40; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
DB 1 WOSAFTK 7

RESULT 2
US-09-140-719-2
Sequence 2, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUOKA, No. US20010026931A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhitro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-2

Query Match 100.0%; Score 40; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTK 7
111111
DB 1 WOSAFTK 7

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSUROUOKA, No. US20020164711A1uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhitro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-MAR-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 40; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
DB 181 WSAFTK 187

RESULT 4

US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, No. US20010026931A1no
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1nhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MECHANOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 40; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFTK 7
DB 181 WSAFTK 187

RESULT 5

US-09-738-626-5215
Sequence 5215, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAKI, AKIO
TITLE OF INVENTION: NOVEL POLYPEPTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentin ver. 3.0
SEQ ID NO 5215
LENGTH: 125
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5215

Query Match 77.5%; Score 31; DB 9; Length 125;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WSAFTK 7
DB 30 WSAFTK 36

RESULT 6

US-09-764-853-656
Sequence 656, Application US/09764853
Patent No. US20020090672A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P3206
CURRENT APPLICATION NUMBER: US/09/764,853
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 656
LENGTH: 168
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-853-656

Query Match 77.5%; Score 31; DB 10; Length 168;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 49 WQSVENK 55

RESULT 7
US-10-091-438-157
Sequence 157, Application US/10091438
Publication No. US20030077606A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT217C1
CURRENT APPLICATION NUMBER: US/10/091,438
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/764,879
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299

PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-06
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 77.5% Score 31: DB 9: Length 194:

Best Local Similarity 71.4% Pred. No. 1e+02: 2: Indels 0: Gaps 0:

Matches 5: Conservative 0: Mismatches

QY 1 WSAFTK 7
||| |

Db 166 WQSVFNK 172
RESULT 8
US-10-007-693-65
Sequence 65, Application US/10007693
Patent No. US20020146776A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
FILE REFERENCE: 210121.515C2
CURRENT APPLICATION NUMBER: US/10/007,693
NUMBER OF SEQ ID NOS: 157
SEQ ID NO 65
LENGTH: 978
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-10-007-693-65

Query Match 77.5% Score 31: DB 12: Length 978:
Best Local Similarity 71.4% Pred. No. 5.2e+02:
Matches 5: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

QY 1 WSAFTK 7
||| |
Db 854 WQSKFTE 860

RESULT 9
US-10-097-065-172
Sequence 172, Application US/10097065
Publication No. US20030055236A1
GENERAL INFORMATION:
APPLICANT: Moore, Paul A. et al.
TITLE OF INVENTION: 110 Human Secreted Proteins
FILE REFERENCE: P2021P1
CURRENT APPLICATION NUMBER: US/10/097,065
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: PCT/US98/27059
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,007
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,057
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,006
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,369
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,367
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,368
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,169
PRIOR FILING DATE: 1997-12-19
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,008
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,365
PRIOR FILING DATE: 1997-12-19
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 172
LENGTH: 195

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (195)
OTHER INFORMATION: Xaa equals stop translation
US-10-097-065-172

Query Match 75.0%; Score 30; DB 9; Length 195;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WQSAFT 7
11111
DB 178 WMSWFTK 184

RESULT 10
US-10-225-567A-516
Sequence 516, Application US/10225567A
Publication No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: Lifespan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burmer, Glenna C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1
SEQ ID NO 516
LENGTH: 385
TYPE: PRT
ORGANISM: Homo sapiens
US-10-225-567A-516

Query Match 75.0%; Score 30; DB 9; Length 385;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WQSAFT 6
11111
DB 241 WQPAFT 246

RESULT 11
US-09-738-626-6127
Sequence 6127, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6127
LENGTH: 687
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6127

Query Match 75.0%; Score 30; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAF 5
11111
DB 377 WQSAF 381

RESULT 12
US-10-204-887-158
Sequence 158, Application US/10204887
Publication No. US20030124569A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: PANZER, Scott R.
APPLICANT: SPIRO, Peter A.
APPLICANT: BANVILLE, Steven C.
APPLICANT: SHAH, Purvi
APPLICANT: CHALUP, Michael S.
APPLICANT: CHANG, Simon C.
APPLICANT: CHEN, Alice
APPLICANT: D'SA, Steven A.
APPLICANT: AMSHEY, Stefan
APPLICANT: DAHL, Christopher R.
APPLICANT: DAM, Tam C.
APPLICANT: DANIELS, Susan E.
APPLICANT: DUFOR, Gerard E.
APPLICANT: FLORES, Vincent
APPLICANT: FONG, Willy T.
APPLICANT: GREENMALT, Lila B.
APPLICANT: HITLMAN, Jennifer L.
APPLICANT: JONES, Anissa L.
APPLICANT: LIU, Tommy F.
APPLICANT: ROSEBERRY, Ann M.
APPLICANT: ROSEN, Bruce H.
APPLICANT: RUSSO, Frank D.
APPLICANT: STOCKREHER, Theresa K.
APPLICANT: DAFEO, Abel
APPLICANT: WRIGHT, Rachel J.
APPLICANT: VAP, Pierre E.
APPLICANT: YU, Jimmy Y.
APPLICANT: BRADLEY, Diana L.
APPLICANT: BRATCHER, Shawn R.
APPLICANT: CHEN, Wensheng
APPLICANT: COHEN, Howard J.
APPLICANT: HODGSON, David M.
APPLICANT: LINCOLN, Stephen E.
TITLE OF INVENTION: SECRETORY MOLECULES
FILE REFERENCE: PT-1134 PCT
CURRENT APPLICATION NUMBER: US/10/204,887
CURRENT FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/185,215; 60/185,216; 60/205,232; 60/205,323; 60/205,2
60/205,324; 60/205,286
PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17
NUMBER OF SEQ ID NOS: 159
SOFTWARE: PERL Program
SEQ ID NO 158
LENGTH: 145
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No. US20030124569A1 LI:221836.3.orf2:2000WAY01
US-10-204-887-158

Query Match 72.5%; Score 29; DB 9; Length 145;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
111111
DB 111 WQSEFK 117

RESULT 13

US-09-738-626-6151
Sequence 6151, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: TOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738.626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6151
LENGTH: 165
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6151

Query Match 72.5%; Score 29; DB 9; Length 165;
Best Local Similarity 57.1%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
111111
DB 80 WDSFTR 86

RESULT 14

US-09-738-626-4605
Sequence 4605, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: TOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738.626
CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4605
LENGTH: 225
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4605

Query Match 72.5%; Score 29; DB 9; Length 225;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
111111
DB 187 WQSAFSE 193

RESULT 15
US-09-854-133-397
Sequence 397, Application US/09854133
Publication No. US20020183499A1
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamath, Radoch
APPLICANT: Henderson, Robert A.
APPLICANT: Benson, Darin R.
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.475C10
CURRENT APPLICATION NUMBER: US/09/854.133
CURRENT FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 735
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 397
LENGTH: 307
TYPE: PRT
ORGANISM: Homo sapiens
US-09-854-133-397

Query Match 72.5%; Score 29; DB 9; Length 307;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WQSAFT 6
111111
DB 212 WQVAF 217

Search completed: July 11, 2003, 12:37:32
Job time : 3.05198 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 1.41164 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-2

Perfect score: 40

Sequence: 1 WQSAFTK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	82.5	397	2 F83721	chromate transport
2	33	82.5	865	2 T40288	hypothetical prote
3	33	82.5	880	2 B45956	beta-glucosidase (
4	32	80.0	77	2 S04795	mobilization prote
5	32	80.0	736	2 C69451	cationic amino aci
6	32	80.0	837	2 AD0159	outer membrane ush
7	32	80.0	938	2 F86548	polymorphic outer
8	32	80.0	938	2 H72074	polymorphic membra
9	31	77.5	112	2 H82752	hypothetical prote
10	31	77.5	289	2 H97480	hypothetical prote
11	31	77.5	289	2 AH2698	permease (imported
12	31	77.5	291	2 T64069	hypothetical prote
13	31	77.5	357	2 G81185	hypothetical prote
14	31	77.5	357	2 F81916	probable integral
15	31	77.5	427	2 T32652	hypothetical prote
16	31	77.5	514	2 E75142	endoglucanase PA80
17	31	77.5	590	2 F95853	probable phospholi
18	31	77.5	677	2 S57599	FUS2 protein - yea
19	31	77.5	722	2 B86583	transcription elon
20	31	77.5	722	2 C72040	transcription elon
21	31	77.5	978	2 G72076	polymorphic membra
22	31	77.5	978	2 B81593	polymorphic membra
23	31	77.5	978	2 C86547	polymorphic outer
24	31	77.5	1108	2 A59251	DNA polymerase III
25	31	77.5	1145	2 A59251	myosin - Acetabula
26	30	75.0	128	2 H97098	probable cation ef
27	30	75.0	165	2 D69784	hypothetical prote
28	30	75.0	261	2 T16550	hypothetical prote
29	30	75.0	266	2 T16106	hypothetical prote

30	30	75.0	273	2 AF2025	hypothetical prote
31	30	75.0	275	2 S75994	hypothetical prote
32	30	75.0	296	2 S72601	hypothetical prote
33	30	75.0	317	2 D84098	hypothetical prote
34	30	75.0	324	2 C95864	conserved hypotnet
35	30	75.0	337	2 G97742	cytochrome c oxida
36	30	75.0	351	1 XUBPB4	DNA beta-glucosylt
37	30	75.0	359	2 D88940	protein C05E4.1 (1
38	30	75.0	415	2 AC0301	probable membrane
39	30	75.0	429	1 AUPB0D	glutamate-amonia
40	30	75.0	432	2 T14292	glutamate-amonia
41	30	75.0	479	2 T20700	hypothetical prote
42	30	75.0	494	2 E70352	anthranilate synth
43	30	75.0	504	2 S75597	hypothetical prote
44	30	75.0	521	2 T49355	related to protein
45	30	75.0	550	2 T40379	hypothetical prote

ALIGNMENTS

```
RESULT 1
F83721
Chromate transporter Chra [Imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F83721
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: F83721
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <STO>
A:Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BAR04293.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: chra
C:Superfamily: chromate resistance protein A

Query Match
Best Local Similarity 82.5%; Score 33; DB 2; Length 397;
Matches 5; Conservativity 71.4%; Pred. No. 28;
Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 160 WQTAFQ 166

RESULT 2
T40288
Hypothetical protein SPBC354.08c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40288
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hildbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21918
A:Accession: T40288
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-865 <MOO>
A:Cross-references: EMBL:AL022071; PIDN:CA117808.1; GSPDB:GN000067; SPDB:SPBC354.08c
A:Experimental source: strain 972h-; cosmid c354
C:Genetics:
A:Gene: SPDB:SPBC354.08c
A:Map position: 2
A:Introns: 58/2

Query Match
Best Local Similarity 82.5%; Score 33; DB 2; Length 865;
Matches 5; Conservativity 71.4%; Pred. No. 65;
Mismatches 1; Indels 0; Gaps 0;
```

OY 1 WOSAFTR 7
||:|
448 WOSAFTR 454

RESULT 3

B43956

beta-glucosidase (EC 3.2.1.21) 2 precursor - yeast (Saccharomycopsis fibuligera)

C:Species: Saccharomycopsis fibuligera

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 24-Sep-1999

C:Accession: B43956

R:Mechida, M.; Ohtsuki, I.; Fukui, S.; Yamashita, I.

Appl. Environ. Microbiol. 54, 3147-3155, 1988

A:Title: Nucleotide sequences of Saccharomycopsis fibuligera genes for extracellular beta-glucosidase

A:Reference number: A45956; MUID:89133518; PMID:3146949

A:Accession: B45956

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-880 <MAC>

A:Cross-references: GB:M22476; NID:g170809; PIDN:AAA34315.1; PID:g170810

C:Superfamily: beta-glucosidase 1

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 82.5%; Score 33; DB 2; Length 880;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTR 7
||:|
56 WOSAFTR 62

RESULT 4

S04795

mobilization protein mobD - Escherichia coli plasmid ColK

C:Species: Escherichia coli

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 18-Jun-1993

C:Accession: S04795

R:Boyd, A.C.; Archer, J.A.K.; Sherratt, D.J.

Mol. Gen. Genet. 217, 488-498, 1989

A:Title: Characterization of the ColE1 mobilization region and its protein products.

A:Reference number: JQ0389; MUID:89364735; PMID:2671664

A:Accession: S04795

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-77 <BOY>

C:Genetics:

A:Gene: mobD; mbxD

A:Genome: plasmid

Query Match 80.0%; Score 32; DB 2; Length 77;
Best Local Similarity 83.3%; Pred. No. 7.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTR 6
||:|
26 WOSAFTR 31

RESULT 5

C69451

cationic amino acid transporter (cat-1) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: C69451

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

, Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: uterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kalne, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: C69451
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-736 <KLE>
A:Cross-references: GB:AE000991; GB:AE000782; NID:g2689314; PIDN:AMB8635.1; PID:g264

Query Match 80.0%; Score 32; DB 2; Length 736;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTR 6
||:|
186 WOSAFTR 191

RESULT 6

AD0159

outer membrane usher protein Psac precursor psac [imported] - Yersinia pestis (strain

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001

C:Accession: AD0159

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G

ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barr

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0159

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-837 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90135.1; PID:g15979355; GSPDB:GN00175

C:Genetics:

A:Gene: psac

C:Superfamily: outer membrane usher protein flmD

Query Match 80.0%; Score 32; DB 2; Length 837;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTR 6
||:|
220 WOSAFTR 225

RESULT 7

F86548

polymorphic outer membrane protein E family [imported] - Chlamydia pneumoniae (st

C:Species: Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: F86548

R:Shirai, M.; Hiraoka, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: AB6491; MUID:20330349; PMID:10873362

A:Accession: F86548

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-938 <STO>

A:Cross-references: GB:BA000008; NID:g8978836; PIDN:BAA98772.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: pmp_15

Query Match 80.0%; Score 32; DB 2; Length 938;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAFTR 6
||:|
642 WOSAFTR 647

RESULT 8
H72074

polymorphic membrane protein E/F family CP0286 [imported] - Chlamydomonas pneumoniae (str)
 C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: H72074; E81593
 R:Kelman, S.; Mitchell, W.; Marathe, R.; Lamell, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: H72074
 A:Molecule type: DNA
 A:Residues: 1-938 <ARN>
 A:Cross-references: GB:AE001631; GB:AE001363; NID:94376750; PIDN:AD18608.1; PID:9437675
 A:Experimental source: strain CML029
 R:Read, T.D.; Brunham, R.C.; Shen, C.; GILL, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: E81593
 A:Molecule type: DNA
 A:Residues: 1-938 <REA>
 A:Cross-references: GB:AE002190; GB:AE002161; NID:97189209; PIDN:AAF38143.1; PID:9718921
 A:Experimental source: strain AR39, HL cells
 C:Genetics:
 A:Gene: pmp_15; CP0286

Query Match 80.0%; Score 32; DB 2; Length 938;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
 111111
 Db 642 WSAFT 647

RESULT 9

H82752
 hypothetical protein XF0867 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: H82752
 R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: H82752
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-112 <SIM>
 A:Cross-references: GB:AE003926; GB:AE003849; NID:99105771; PIDN:AAF83677.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Biondo, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrer, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 Submitted to Genbank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0867

Query Match 77.5%; Score 31; DB 2; Length 112;

Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
 111111
 Db 95 WSAFT 100

RESULT 10

H97480
 hypothetical protein AGR_C_1820 [imported] - Agrobacterium tumefaciens (strain C58, C
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: H97480
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:11743194
 A:Accession: H97480
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86801.1; PID:915156007; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_1820
 A:Map position: circular chromosome

Query Match 77.5%; Score 31; DB 2; Length 289;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
 111111
 Db 41 WSAFT 46

RESULT 11

AH2698
 permealase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AH2698
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McCl
 erage, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AH2698
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <KUR>
 A:Cross-references: GB:AE006688; PIDN:AAL42006.1; PID:917739380; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0992
 A:Map position: circular chromosome

Query Match 77.5%; Score 31; DB 2; Length 289;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
 111111
 Db 41 WSAFT 46

RESULT 12

hypothetical protein HI0461 - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae
 C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
 C:Accession: 164069
 R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodet, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M. Science 265, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: 164069
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-291 <TIGR>
 A:Cross-references: GB:U32728; GB:L42023; NID:g1573425; PIDN:AAC22119.1; PID:g1573435; T C:Superfamily: conserved hypothetical protein H10461

Query Match 77.5% Score 31; DB 2; Length 291;
 Best Local Similarity 71.4% Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTR 7
 ||| |||
 Db 58 WOSGFTQ 64

RESULT 13
 G81185
 hypothetical protein NMB0552 [imported] - Neisseria meningitidis (strain MC58 serogroup C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: G81185
 R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizze, M. Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, A.; Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: G81185
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-357 <PRT>
 A:Cross-references: GB:AE002411; GB:AE002098; NID:g7225776; PIDN:AAF40981.1; PID:g722577 A:Experimental source: serogroup B, strain MC58
 C:Genetics:

Query Match 77.5% Score 31; DB 2; Length 357;
 Best Local Similarity 71.4% Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WOSAFTR 7
 ||| |||
 Db 286 WWSGFTK 292

RESULT 14
 F81916
 probable integral membrane protein NMA0732 [imported] - Neisseria meningitidis (strain Z C:Species: Neisseria meningitidis
 C>Date: 03-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: F81916
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtz, S.; Jorgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: F81916
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-357 <PAR>
 A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84016.1; PID:g737945

A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA0732

Query Match 77.5% Score 31; DB 2; Length 357;
 Best Local Similarity 71.4% Pred. No. 67;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WOSAFTR 7
 ||| |||
 Db 286 WWSGFTK 292

RESULT 15
 T32652
 hypothetical protein F39C12.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T32652
 R:Chisoe, S.; Sansone, J. submitted to the EMBL Data Library, December 1997
 A:Description: The sequence of C. elegans cosmid F39C12.
 A:Reference number: Z21206
 A:Accession: T32652
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-427 <CH1>
 A:Cross-references: EMBL:AF039043; PIDN:AA894196.1; GSPDB:GN00028; CESP:F39C12.3
 A:Experimental source: strain Bristol N2; clone F39C12
 C:Genetics:
 A:Gene: CESP:F39C12.3
 A:Map position: X
 A:Introns: 42/3; 104/3; 133/3; 164/3; 213/3; 276/3; 336/3

Query Match 77.5% Score 31; DB 2; Length 427;
 Best Local Similarity 71.4% Pred. No. 81;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WOSAFTR 7
 ||| |||
 Db 230 WQNALTR 236

Search completed: July 11, 2003, 12:00:34
 Job time: 4.51164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 0.669439 Seconds

(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-2

Sequence: 40
1 MOSAFTRK 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	380	1 SPB7_HUMAN	O75635 homo sapien
2	33	82.5	880	1 BGL2_SACFI	P22507 saccharomyc
3	32	80.0	369	1 H183_RHIL0	O98B00 rhizobium l
4	32	80.0	837	1 PSAC_YERPE	P31527 yersinia pe
5	32	80.0	837	1 PSAC_YERPS	O56983 yersinia ps
6	32	80.0	938	1 PM15_CHLPN	O92883 chlamydia p
7	32	80.0	959	1 G2D1_HUMAN	O9uh19 h general t
8	31	77.5	291	1 Y461_HAEIN	O57144 haemophilus
9	31	77.5	677	1 FUS2_YEAST	O05670 saccharomyc
10	31	77.5	722	1 GRE4_CHLPN	O92794 chlamydia p
11	31	77.5	978	1 PM14_CHLPN	O92895 chlamydia p
12	30	75.0	296	1 UPBS_MYCLE	P38119 mycobacteri
13	30	75.0	351	1 GSTB_BPT4	P04547 bacteriopho
14	30	75.0	385	1 PAR4_HUMAN	O96r10 homo sapien
15	30	75.0	429	1 GLN4_PHAVU	P15102 phaseolus v
16	30	75.0	432	1 GLN2_DAVCA	O22506 dancus caro
17	30	75.0	494	1 TRPE_AQUAE	O66849 aquifex aeo
18	30	75.0	653	1 YDCP_ECOLI	P76104 escherichia
19	30	75.0	917	1 GCP3_DROME	O94yp8 drosophil
20	30	75.0	1210	1 YEH1_ECOLI	P33346 escherichia
21	29	72.5	137	1 V44_SOLIN	P35777 soleosopsis
22	29	72.5	225	1 Y991_COREL	P38118 corynebacte
23	29	72.5	259	1 Y714_STRFR	P29182 streptomyc
24	29	72.5	262	1 Y486_MYCTU	O53434 mycobacteri
25	29	72.5	270	1 ALYS_BPRIT	O38135 bacteriopho
26	29	72.5	345	1 HEMS_YEREN	P31517 yersinia en
27	29	72.5	345	1 HMUS_YERPE	O56990 yersinia pe
28	29	72.5	351	1 UXUA_CLOAB	O971f4 clostridium
29	29	72.5	359	1 MOAL_MYCTU	O05786 mycobacteri
30	29	72.5	405	1 IFSP_MOUSE	P70438 mus musculu
31	29	72.5	450	1 DGT1_PASWU	O9cmi1 pasteurella
32	29	72.5	481	1 GIGA_RHIL0	O985p2 rhizobium l
33	29	72.5	580	1 FZ0B_XENLA	O9w742 xenopus lae

34	29	72.5	586	1 FZ0A_XENLA	O9den5 xenopus lae
35	29	72.5	601	1 YE06_SCHPO	O13803 schizosacch
36	29	72.5	639	1 V70K_PLRPI	P17519 potaro leaf
37	29	72.5	639	1 V70K_PLRPI	P11622 potaro leaf
38	29	72.5	639	1 V4T1_PIRAB	O9uxu2 pyrococcus
39	29	72.5	693	1 TGM3_HUMAN	O08188 homo sapien
40	29	72.5	714	1 GRE4_CHLMU	O9plu1 chlamydia m
41	29	72.5	787	1 YD5A_SCHPO	O10313 schizosacch
42	29	72.5	1124	1 PHY1_TOBAC	P33530 nicotiana t
43	29	72.5	1935	1 YBBA_SCHPO	O60072 schizosacch
44	29	72.5	2376	1 YIM9_YEAST	P40468 saccharomyc
45	28	70.0	77	1 MBED_ECOLI	P13660 escherichia

ALIGNMENTS

RESULT 1
SPB7_HUMAN STANDARD: PRT: 380 AA.
ID SPB7_HUMAN
AC O75635:
15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Megsin (TP55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97326116; PubMed-9182567;
RA Tsujimoto M., Tsuruoka N., Ishida N., Kurihara T., Iyasa F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Haruyama M., Miura K.,
RA Nakanishi T., Nakazato H., Teramura M., Mitsuuchi H., Yamaguchi N.,
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity."
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-98376492; PubMed-9710452;
RA Miyata T., Nangaku M., Suzuki D., Inagi R., Uragami K., Sakai H.,
RA Okubo K., Kurokawa K.;
RT "A mesangium-predominant gene, megsin, is a new serpin upregulated in
RT Iga nephropathy."
RL J. Clin. Invest. 102:828-836(1998).
CC -I- FUNCTION: Might function as an inhibitor of lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -I- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL: D88575; BAA3132.1; -
CC EMBL: AF027866; AAC64506.1; -
CC HSSP: P05619; 1HDE.
CC Genew: HGNC:13902; SERPINB7.
CC MIM: 603357; -
CC InterPro: IPR000215; Serpin.
CC Pfam: PF00079; serpin; 1.
CC SMART: SM00093; SERPIN; 1.
CC PROSITE: PS00284; SERPIN; 1.

KM Serpin; Serine protease inhibitor.
 FT ACT_SITE 347 348 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA: 42904 MW: 9A2CDB6C63CFF605 CRC64:

Query Match 100.0%; Score 40; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.36;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOSAFK 7
 DB 181 MOSAFK 187

RESULT 2
 BGL2_SACFI STANDARD; PRT; 880 AA.
 ID BGL2_SACFI
 AC P22507;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-glucosidase 2 precursor (EC 3.2.1.21) (Gentloblase) (Cellloblase)
 DE (beta-D-glucoside glucohydrolase).
 GN BGL2.
 OS Saccharomycopsis fibuligera (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
 OX NCBI_TaxID:4944;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89133518; PubMed-3146949;
 RA Machida M., Ohtsuki I., Fukui S., Yamashita I.;
 RT "Nucleotide sequences of Saccharomycopsis fibuligera genes for
 RT extracellular beta-glucosidases as expressed in Saccharomyces
 RT cerevisiae";
 RT Appl. Environ. Microbiol. 54:3147-3155(1988).
 RL -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- PATHWAY: Cellulose degradation.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M22476; AAA34315.1; -;
 DR InterPro: IPR002772; GH_3C.
 DR InterPro: IPR001764; GH_3N.
 DR Pfam: PF00933; Glyco_hydro_3_1.
 DR Pfam: PF01915; Glyco_hydro_3_C_1.
 DR PRINTS: PR00133; GLHDLASE3.
 DR PROSITE: PS00775; GLYCOSYL-HYDROL_F3_1.
 DR PROSITE: PS00775; GLYCOSYL-HYDROL_F3_1.
 KW Hydrolyase; Glycosidase; Cellulose degradation; Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 880
 FT ACT_SITE 299 299
 FT CARBOHYD 24 24
 FT CARBOHYD 77 77
 FT CARBOHYD 271 271
 FT CARBOHYD 336 336
 FT CARBOHYD 343 343
 FT CARBOHYD 376 376
 FT CARBOHYD 548 548
 FT CARBOHYD 589 589
 FT CARBOHYD 712 712
 FT CARBOHYD 743 743
 FT CARBOHYD 794 794
 SQ SEQUENCE 880 AA: 96799 MW: C95061283090814C CRC64;

Query Match 82.5%; Score 33; DB 1; Length 880;
 Best Local Similarity 71.4%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOSAFK 7
 DB 56 WODATK 62

RESULT 3
 H183_RHIL0 STANDARD; PRT; 369 AA.
 ID H183_RHIL0
 AC Q98B00;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Histidinol phosphate aminotransferase 3 (EC 2.6.1.9) (Imidazole
 DE acetyl-phosphate transaminase 3).
 GN H183 OR M185786.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Mesorhizobium.
 OX NCBI_TaxID:381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-MAF303099;
 RX MEDLINE-21082930; PubMed-11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti";
 RT Mesorhizobium loti";
 RL DNA Res. 7:331-338(2000).
 CC -1- CATALYTIC ACTIVITY: L-histidinol-phosphate + 2-oxoglutarate = 3-
 CC (imidazol-4-yl)-2-oxopropyl phosphate + L-glutamate.
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -1- PATHWAY: Histidine biosynthesis; seventh step.
 CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES. HISTIDINOL-PHOSPHATE AMINOTRANSFERASE
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AP003007; BAB52172.1; -;
 DR InterPro: IPR004839; AminoTransf1/2.
 DR InterPro: IPR001917; NHTransf2.
 DR Pfam: PF00155; aminoTransf_1_2_1.
 DR TIGRFAMs: TIGR01141; h18c; 1.
 DR PROSITE: PS00599; AA_TRANSFER_CLASS_2; 1.
 DR PROSITE: PS00599; AA_TRANSFER_CLASS_2; 1.
 KW Histidine biosynthesis; Transferase; Aminotransferase;
 KW Pyridoxal phosphate; Complete proteome.
 FT BINDING 220 220
 FT SEQUENCE 369 AA: 40710 MW: DCB33E0507165570 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 369;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MOSAFK 7
 DB 315 MOSAFK 321

RESULT 4
 PSAC_YERPE

ID PSAC_YERPE STANDARD: PRT: 837 AA.
AC P31537: 056980: 01-JUL-1993 (Rel. 26, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Outer membrane usher protein psc precursor.
GN PSAC OR YPO1305.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
NCBI_TaxID=632.
OX NCBI_TaxID=632.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-EV76:
RA Cherepanov P.A.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-CO-92 / Blovat Orientalis;
MEDLINE-21470413: PubMed-11586360;
RA Parthall J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
Prentice M.B., Sepaldia M., James K.D., Churche C., Mungall K.L.,
Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
Fellwell T., Hamlin N., Holroyd S., Jagsis K., Karlyshev A.V.,
Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN (3)
RP SEQUENCE OF 1-196 FROM N.A.
RX MEDLINE-93302507: PubMed-8100346;
RA Lindler L.E., Tall B.D.;
RT "Yersinia pestis PH 6 antigen forms fimbriae and is induced by
intercellular association with macrophages.";
RL Microbiol. 8:311-324(1993).
CC -1- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF PSA (PH 6)
FIMBRIAL SUBUNITS ACROSS THE OUTER MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
(By similarity).
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X97759: CAA66357.1: ALT_INIT.
DR EMBL: A414147: CAC90135.1: -;
DR EMBL: M87713: AAA27664.1: ALT_INIT.
DR PIR: S27749: S27749.
DR PIR: S32928: S32928.
DR InterPro: IPR000015: Fimb_usher.
DR Pfam: PF00577: usher.1.
DR PROSITE: PS01151: FIMBRIAL_USHER.1.
KW Outer membrane; Transmembrane; Fimbria; Transport; Signal;
KW Complete proteome.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 837 OUTER MEMBRANE USHER PROTEIN PSAC.
FT CHAIN 172 177 ANISER -> EISYK (IN REF. 1 AND 3).
FT CONFLICT 178 196 ORONGYRSGTITLYPG -> SFVKMVDGVHNTSYIP
V (IN REF. 3).
FT CONFLICT 395 395 R -> A (IN REF. 1).
FT CONFLICT 437 437 E -> K (IN REF. 1).
FT CONFLICT 797 805 ETGOLVQV -> KQSGYCM (IN REF. 1).
SO SEQUENCE 837 AA: 91702 MW: 91702 MW: D901C236C012B2A CRC64;

Query Match 80.0%; Score 32; DB 1; Length 837;
Best Local Similarity 83.3%; Pred. NO. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAF 6
Db 220 WOSAF 225

RESULT 5
PSAC_YERPS
ID PSAC_YERPS STANDARD: PRT: 837 AA.
AC Q56963;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Outer membrane usher protein psc precursor.
GN PSAC.
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
NCBI_TaxID=633.
OX NCBI_TaxID=633.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-YPIII;
RX MEDLINE-96294755: PubMed-8698470;
RA Yang Y., Merliam J.J., Mueller J., Isberg R.R.;
RT "The psa locus is responsible for thermolabile binding of Yersinia
pseudotuberculosis to cultured cells.";
RL Infect. Immun. 64:2483-2489(1996).
CC -1- FUNCTION: INVOLVED IN THE EXPORT AND ASSEMBLY OF PSA (PH 6)
FIMBRIAL SUBUNITS ACROSS THE OUTER MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
(By similarity).
CC -1- SIMILARITY: BELONGS TO THE FIMBRIAL EXPORT USHER FAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L76301: AAC37058.1: ALT_INIT.
DR InterPro: IPR000015: Fimb_usher.
DR Pfam: PF00577: usher.2.
DR PROSITE: PS01151: FIMBRIAL_USHER.1.
KW Outer membrane; Transmembrane; Fimbria; Transport; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 837 OUTER MEMBRANE USHER PROTEIN PSAC.
SO SEQUENCE 837 AA: 91786 MW: 820F7896DE4BEA1A CRC64;

Query Match 80.0%; Score 32; DB 1; Length 837;
Best Local Similarity 83.3%; Pred. NO. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WOSAF 6
Db 220 WOSAF 225

RESULT 6
PM15_CHLPN
ID PM15_CHLPN STANDARD: PRT: 938 AA.
AC Q92883;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp15 precursor (Polymorphic membrane
protein 15).
GN PMP15 OR CPN0466 OR CP0286.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN (1)

RP SEQUENCE FROM N.A.
 RC STRAIN-CWL029;
 RX MEDLINE-99206606; PubMed-10192388;
 RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AR39;
 RX MEDLINE-20150255; PubMed-10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Galun M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Morn and Chlamydia
 RT pneumoniae AR39";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J138;
 RX MEDLINE-20330349; PubMed-10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CML029 from USA";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 CC (POTENTIAL)
 CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: AE001631; AAD18608.1; -
 CC EMBL: AE002190; AAF38143.1; -
 CC EMBL: AP002546; BAA98672.1; -
 CC TRNG: CP0286; -
 CC DR InterPro: IPR003368; Chlamydia_PMP.
 CC DR InterPro: IPR003357; OMP.
 CC DR Pfam: PF02385; OMP. 1.
 CC DR Pfam: PF02415; DUF145; 1.
 CC KW Outer membrane, Signal, Multigene family, Complete proteome.
 CC FT SIGNAL 17 POTENTIAL.
 CC FT CHAIN 18 938 PROBABLE OUTER MEMBRANE PROTEIN PMP15.
 CC SQ SEQUENCE 938 AA; 102194 MW; A904AB05B567455 CRC64;
 CC
 CC Query Match 80.0%; Score 32; DB 1; Length 938;
 CC Best Local Similarity 83.3%; Pred. NO. 45;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 WQSAYT 6
 CC Db 642 WQSAYT 647
 CC
 CC RESULT 7
 CC ID G2D1_HUMAN STANDARD: PRT; 959 AA.
 CC AC 09UHL9: 095444; 09UHL8: 09U191; 08WVC4;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DE 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE General transcription factor II-I repeat domain-containing protein 1
 CC (GTF2I repeat domain containing protein 1) (Muscle TFI-I repeat
 CC domain-containing protein 1) (General transcription factor I1) (Slow-
 CC muscle-fiber enhancer binding protein) (USE BI binding protein)

DE (MUSTRD1/BEN) (Williams-Beuren syndrome chromosome region 11 protein).
 GN GTF2IRD1 OR MUSTRD1 OR GTF3 OR CREM1 OR RBAP2 OR WBSCTRL1 OR WBSCTRL2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCBI_TaxId:9606;
 CC
 CC [1]
 CC SEQUENCE FROM N.A. (ISOFORM 2).
 CC TISSUE=Muscle;
 CC MEDLINE-98449952; PubMed-9774679;
 CC O'Mahoney J.V., Guven K.L., Lin J., Joya J.E., Robinson C.S.,
 CC Wade R.P., Hardeman E.C.;
 CC "Identification of a novel slow-muscle-fiber enhancer binding protein,
 CC MUSTRD1";
 CC Mol. Cell. Biol. 18:6641-6652(1998).
 CC [2]
 CC SEQUENCE FROM N.A. (ISOFORM 2).
 CC TISSUE=Colon carcinoma;
 CC MEDLINE-99216421; PubMed-10198167;
 CC Osborne L.R., Campbell T., Daradich A., Scherer S.W., Tsui L.-C.;
 CC "Identification of a putative transcription factor gene (WBSCTRL1) that
 CC is commonly deleted in Williams-Beuren syndrome";
 CC Genomics 57:279-284(1999).
 CC [3]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC TISSUE=Fetal brain;
 CC MEDLINE-20037629; PubMed-10573005;
 CC Tassabehji M., Carette M., Wilmet C., Donnai D., Read A.P.,
 CC Metcalfe K.;
 CC "A transcription factor involved in skeletal muscle gene expression is
 CC deleted in patients with Williams syndrome";
 CC Eur. J. Hum. Genet. 7:737-747(1999).
 CC [4]
 CC SEQUENCE FROM N.A. (ISOFORM 2).
 CC MEDLINE-20044629; PubMed-10575229;
 CC Franke Y., Peoples R.J., Francke U.;
 CC "Identification of GTF2IRD1, a putative transcription factor within
 CC the Williams-Beuren syndrome deletion at 7q11.23";
 CC Cytogenet. Cell Genet. 86:296-304(1999).
 CC [5]
 CC SEQUENCE FROM N.A. (ISOFORM 1), INTERACTION WITH RB1, AND MUTAGENESIS.
 CC TISSUE=Cervical carcinoma, placenta, and fetal spleen;
 CC MEDLINE-20115113; PubMed-10642537;
 CC Yan X., Zhao X., Qian M., Guo N., Gong X., Zhu X.;
 CC "Characterization and gene structure of a novel retinoblastoma-
 CC protein-associated protein similar to the transcription regulator
 CC TFI-I";
 CC Biochem. J. 345:749-757(2000).
 CC [6]
 CC SEQUENCE FROM N.A. (ISOFORM 1).
 CC TISSUE=Uterus;
 CC Strausberg R.;
 CC Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC [7]
 CC FUNCTION.
 CC MEDLINE-21332325; PubMed-11438732;
 CC Tussie-Luna M.I., Bayarsaihan D., Ruddle F.H., Roy A.L.;
 CC "Repression of TFI-I-dependent transcription by nuclear exclusion";
 CC Proc. Natl. Acad. Sci. U.S.A. 98:7789-7794(2001).
 CC -1- FUNCTION: May be a transcription regulator involved in cell-cycle
 CC progression and skeletal muscle differentiation. May repress GTF2I
 CC transcriptional functions, by preventing its nuclear residency, or
 CC by inhibiting its transcriptional activation. May contribute to
 CC slow-twitch fiber type specificity during myogenesis and in
 CC regenerating muscles. Binds tropomyosin I slow-muscle fiber enhancer
 CC (USE BI). Binds specifically and with high affinity to the EFG
 CC sequences derived from the early enhancer of HOXC8 (By
 CC similarity).
 CC -1- SUBUNIT: Interacts with the retinoblastoma protein (RB1) via its
 CC C-terminus.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2; are
 CC produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: Highly expressed in adult skeletal muscle,
 CC heart, fibroblast, bone and fetal tissues. Expressed at lower
 CC levels in all other tissues tested.
 CC -1- DEVELOPMENTAL STAGE: Highly expressed in developing and
 CC regenerating muscles, at the time of myofiber diversification.
 CC -1- DOMAIN: The N-terminal half may have an activating activity.
 CC -1- DISEASE: Haploinsufficiency of GTF2I1RD1 may be the cause of
 CC certain cardiovascular and musculo-skeletal abnormalities observed
 CC in Williams-Beuren syndrome (WBS), a rare developmental disorder.
 CC It is a contiguous gene deletion syndrome involving genes from
 CC chromosome band 7q11.23.
 CC -1- SIMILARITY: BELONGS TO THE TFI-I FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 GTF2I-LIKE REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF118270; AAD14687.2; -
 CC DR EMBL; AF104923; AAD27668.1; -
 CC DR EMBL; AF151354; AAF19786.1; -
 CC DR EMBL; AF156489; AAF17358.1; -
 CC DR EMBL; AF089107; AAF21796.1; -
 CC DR EMBL; BC018136; AAH18136.1; -
 CC DR Gene; HGNC:4661; GTF2I1RD1.
 CC DR MIM; 604318; -
 CC DR MIM; 194050; -
 CC DR InterPro: IPR004212; GTF2I.
 CC DR Pfam: PF02946; GTF2I; 5.
 CC KW Transcription regulation; Developmental protein; DNA-binding;
 CC KW Nuclear protein; Repeat; Alternative splicing; Polymorphism;
 CC KW Williams-Beuren syndrome.
 CC FT REPEAT 128 203 GTF2I 1.
 CC FT REPEAT 351 426 GTF2I 2.
 CC FT REPEAT 565 640 GTF2I 3.
 CC FT REPEAT 705 780 GTF2I 4.
 CC FT REPEAT 802 877 GTF2I 5.
 CC FT DOMAIN 898 905 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 906 930 SER-RICH.
 CC FT VARSPLIC 656 670 MISSING (IN ISOFORM 2).
 CC FT VARIANT 652 652 M -> V (IN DBSNP:2301895).
 CC FT MUTAGEN 898 959 /FTID-VAR_013446.
 CC FT CONFLICT 111 111 MISSING: CYTOPLASMIC LOCALIZATION.
 CC FT CONFLICT 378 378 G -> S (IN REF. 1 AND 2).
 CC FT CONFLICT 378 378 R -> Q (IN REF. 5).
 CC SQ SEQUENCE 959 AA; 106057 MW; 7DA3097879701540 CRC64;
 CC -----
 CC Query Match 80.0%; Score 32; DB 1; Length 959;
 CC Best Local Similarity 71.4%; Pred. No. 47;
 CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC Oy 1 WQSAFTK 7
 CC | 11111;
 CC Db 20 WNSAFTR 26
 CC -----
 CC RESULT 8
 CC ID Y461_HAEIN STANDARD: PRT; 291 AA.
 CC AC 057144; 005022;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Protein HI0461.
 CC GN HI0461.
 CC OS Haemophilus influenzae.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC OC Haemophilus.
 CC NCBI_TaxID=727;
 CC OX

RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-Rd / KW20 / ATCC 51907;
 RC MEDLINE=95350630; PubMed=7542800;
 RX Felschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavang A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Usterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
 RA Fine L.D., Fitchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RP Rd.";
 RL Science 269:496-512(1995).
 RL [2]
 RN IDENTIFICATION BY MASS SPECTROMETRY.
 RP MEDLINE=20137488; PubMed=10675023;
 RX Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 RA Gray C., Fountoulakis M.;
 RT "Two-dimensional map of the proteome of Haemophilus influenzae.";
 RL Electrophoresis 21:411-429(2000).
 CC -1- SIMILARITY: TO E.COLI YBJX AND WEAK, TO P.HAEMOLYTICA LAMP.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U32728; AAC22119.1; -
 CC DR TIGR; HI0461; -
 CC DR Complete proteome.
 CC SQ SEQUENCE 291 AA; 34846 MW; 699FDC4204A2E2A0 CRC64;
 CC -----
 CC Query Match 77.5%; Score 31; DB 1; Length 291;
 CC Best Local Similarity 71.4%; Pred. No. 22;
 CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC Oy 1 WQSAFTK 7
 CC | 11111;
 CC Db 58 WQSLFTQ 64
 CC -----
 CC RESULT 9
 CC ID FUS2_YEAST STANDARD: PRT; 677 AA.
 CC AC 005670; 005023;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Nuclear fusion protein FUS2.
 CC GN FUS2 OR YMR232W OR YMG959.14.
 CC OS Saccharomyces cerevisiae (baker's yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CC NCBI_TaxID=4932;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-S288C;
 CC RX MEDLINE=96032260; PubMed=7559752;
 CC RA Elion E.A., Trueheart J., Fink G.R.;
 CC RT "Fus2 localizes near the site of cell fusion and is required for both
 CC cell fusion and nuclear alignment during zygote formation.";
 CC J. Cell Biol. 130:1283-1296(1995).
 CC [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-S288C / AB972;
 CC RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 CC Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.


```

CC -1- FUNCTION: PROMOTES CELL FUSION DURING ZYGOTE FORMATION.
CC -1- SUBCELLULAR LOCATION: LOCALIZES NEAR THE SITE OF CELL FUSION.
CC -1- SIMILARITY: CONTAINS 1 DBI-HOMOLOGY (DBI) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X90752; CA62275.1; -.
DR EMBL: Z49939; CA90203.1; -.
DR SGD: S0004845; FUS2.
DR InterPro: IPR001331; GDS_CDC24.
DR InterPro: IPR000219; RHOGEP.
DR SMART: SM00325; RHOGEP; 1.
DR PROSITE: PS00741; DH_1; FALSE_NEG.
DR PROSITE: PS50010; DH_2; 1.
KW FUSION PROTEIN.
FT DOMAIN 112 326 DH.
FT CONFLICT 185 185 L->R (IN REF. 1).
FT CONFLICT 203 203 F->L (IN REF. 1).
FT CONFLICT 543 546 MISSING (IN REF. 1).
FT CONFLICT 589 677 HDTECLINLVIKVFLKYLITLTIAGKKYKQKLDENMSLND
IATGQIKNDILQCYKSRKMTKRMKMDPFPDPPGSGRY
VRKLFEL -> PRYRMVLYELVQSLKIFGNHCKMKKIPA
KRS (IN REF. 1).
SQ SEQUENCE 677 AA; 79036 MW; D7C89ED0738BA31F CRC64;

Query Match
Best Local Similarity 77.5%; Score 31; DB 1; Length 677;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WOSAFK 7
DB 199 WOKIFK 205

RESULT 10
GREA_CHLPN STANDARD; PRT; 722 AA.
AC Q9Z7G4; Q9J0E9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription elongation factor grea (Transcript cleavage factor
DE grea).
GN GREA OR CPN0741 OR CP0004.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uitterlind T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shiba T., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hatori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and W4029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION
CC ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING
CC SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACTION
CC OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN
CC LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY
CC CLEAVAGE FACTORS SUCH AS GREA OR GREB ALLOWS THE RESUMPTION OF
CC ELONGATION FROM THE NEW 3' TERMINUS. GREA RELEASES SEQUENCES OF
CC 2 TO 3 NUCLEOTIDES (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 GREA ASSOCIATED DOMAIN 1 (GRAD1).
CC -1- SIMILARITY: BELONGS TO THE GREA/GREB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF001656; AAD18880.1; -.
DR EMBL: AE002164; AAF37801.1; -.
DR EMBL: AP002547; BAA98948.1; -.
DR TIGR: CP0004; -.
DR InterPro: IPR001437; Grea_Greb.
DR Pfam: PF03449; Grea_Greb_N; 1.
DR Prodom: PD004918; Grea_Greb; 1.
DR PROSITE: PS00829; GREAB_1; 1.
DR PROSITE: PS00830; GREAB_2; 1.
KW Transcription regulation; DNA-binding; Coiled coil; Complete proteome.
FT DOMAIN 1 504 GRAD1.
FT DOMAIN 565 722 GREA.
FT DOMAIN 606 644 COILED COIL (POTENTIAL).
SQ SEQUENCE 722 AA; 82642 MW; E280EDB124D5B8CF CRC64;

Query Match
Best Local Similarity 77.5%; Score 31; DB 1; Length 722;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WOSAFK 7
DB 250 WOSAKTK 256

RESULT 11
PML4_CHLPN STANDARD; PRT; 978 AA.
AC Q9Z895; Q9RB63;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pml4 precursor (Polymorphic membrane
DE protein 14).
GN PMP14 OR CPN0454 OR CP0298.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).

```



```

RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39:
RX MEDLINE-20150255: PubMed-10684935:
RA Read T.D., Brunham R.C., Shen C., Gail S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Baas S.,
RA Linher K., Weidman J., Khouli H., Craven B., Bowman C., Dodson R.,
RA Gwyn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.:
RT "Genome sequences of Chlamydia trachomatis M09 and Chlamydia
RT pneumoniae A39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN-J138:
RX MEDLINE-20330349: PubMed-10871362:
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.:
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001629; AAD18596.1; -
DR EMBL: AE002191; AAF38155.1; -
DR EMBL: AP002546; BAA98661.1; -
DR PHC1-2DPAGE: Q92895; -
DR TIGR: CP0298; -
DR InterPro: IPR003368; Chlamydia_PMP.
DR InterPro: IPR003357; OMP.
DR Pfam: PF02385; OMP; 1.
DR Pfam: PF02415; DUF145; 1.
RW Outer membrane; signal; Multigene family; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 978 PROBABLE OUTER MEMBRANE PROTEIN PMP14.
FT CONFLICT 379 379 N->D (IN REF. 1).
SQ SEQUENCE 978 AA; 103654 MW; 6EEB142999D3019 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 978;
Best Local Similarity 71.4%; Pred. NO. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 854 WQSKFTE 860

RESULT 12
UPPS_MYCLE STANDARD: PRT: 296 AA.
ID UPPS_MYCLE
AC P38119;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) (UPP synthetase)
DE (D1-trans-poly-cis-decaprenyltransferase) (Undecaprenyl diphosphate
DE synthase) (UDS).
GN UPPS OR ML0634 OR B1937_F2_65.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.:
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN-TN:
RX MEDLINE-21128732: PubMed-11234002:
RA Cole S.T., Eigleiner K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsels K., Lacroix C., Maelen J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.:
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: GENERATES UNDECAPRENYL PYROPHOSPHATE (UPP) FROM
CC ISOPRENYL PYROPHOSPHATE (IPP). UPP IS THE PRECURSOR OF THE
CC CARRIER LIPID FOR PEPTIDOGLYCAN SYNTHESIS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: D1-trans-poly-cis-decaprenyl diphosphate +
CC isopentenyl diphosphate -> diphosphate + d1-trans-poly-cis-
CC undecaprenyl diphosphate.
CC -1- SIMILARITY: BELONGS TO THE UPP SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U00016; AA17169.1; -
DR EMBL: AL583919; CAC30142.1; -
DR Leprosoma: ML0634; -
DR InterPro: IPR001441; UPP_synth.
DR Pfam: PF01255; UPP_synthetase; 1.
DR ProDom: PD003461; UPP_synth; 1.
DR TIGRFAMs: TIGR00055; upps; 1.
DR PROSITE: PS01066; UPP_SYNTHETASE; 1.
RW Transferase; Cell division; Cell wall; Peptidoglycan synthesis;
KW Complete proteome.
SQ SEQUENCE 296 AA; 33890 MW; FF4D28B49E7C6545 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 296;
Best Local Similarity 57.1%; Pred. NO. 36;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 257 WQAAVTE 263

RESULT 13
GSTR_BPT4 STANDARD: PRT: 351 AA.
ID GSTR_BPT4
AC P04547; Q38417;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA beta-glucosyltransferase (EC 2.4.1.27) (BGT).
GN BGT OR BETA-GT.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86067181: PubMed-299696;
RA Tomaszewski J., Gram H., Grubb J.W., Rueger W.:
RT "T4-induced alpha- and beta-glucosyltransferase: cloning of the genes

```

RT and a comparison of their products based on sequencing data.";
 RL Nucleic Acids Res. 13:7551-7568(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kutter E., Arita F., Kunisawa T., Tsugita A., Mosig G.,
 RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
 RT "Bacteriophage T4 genome analysis.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-88 FROM N.A.
 RX MEDLINE-88169543; PubMed-2832395;
 RA Thyleen C.;
 RT "Expression and DNA sequence of the cloned bacteriophage T4 dCMP
 RT hydroxymethylase gene.";
 RL J. Bacteriol. 170:1994-1998(1988).
 RN [4]
 RP SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE-87231080; PubMed-3295783;
 RA Lamm N., Tomaszewski J., Rueger W.;
 RT "Nucleotide sequence of the deoxyribidylate hydroxymethylase gene of
 RT bacteriophage T4 (942) and the homology of its gene product with
 RT thymidylate synthase of E. coli.";
 RL Nucleic Acids Res. 15:3920-3920(1987).
 RN [5]
 RP SEQUENCE OF 239-351 FROM N.A.
 RX MEDLINE-92353535; PubMed-1631169;
 RA Sharma M., Ellis R.L., Hinton D.M.;
 RT "Identification of a family of bacteriophage T4 genes encoding
 RT proteins similar to those present in group I introns of fungi and
 RT phage.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6658-6662(1992).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE-94341252; PubMed-8062817;
 RA Veilink A., Rueger W., Driessen H.P.C., Freemont P.S.;
 RT "Crystal structure of the DNA modifying enzyme beta-
 RT glucosyltransferase in the presence and absence of the substrate
 RT uridine diphosphoglucose.";
 RL EMBO J. 13:3413-3422(1994).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE-99428644; PubMed-10497034;
 RA Morea S., Imbercy A., Aschke-Sonnenborn U., Ruger W., Freemont P.S.;
 RT "T4 phage beta-glucosyltransferase: substrate binding and proposed
 RT catalytic mechanism.";
 RL J. Mol. Biol. 292:717-730(1999).
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF GLUCOSE (GLC) FROM URIDINE
 CC DIPHOSPHOGLUCOSE (UDP-GLC) TO 5-HYDROXYMETHYLCYTOSINE (5-HMC) IN
 CC DOUBLE-STRANDED DNA.
 CC -1- CATALYTIC ACTIVITY: Transfers a beta-D-glucosyl residue from UDP-
 CC glucose to an hydroxymethylcytosine residue in DNA.
 CC -1- PATHWAY: INVOLVED IN THE DNA MODIFICATION PROCESS TO PROTECTS
 CC THE PHAGE GENOME AGAINST ITS OWN NUCLEASES AND THE HOST
 CC RESTRICTION ENDONUCLEASE SYSTEM.
 CC -1- SUBUNIT: MONOMER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X03139; CAA26908.1; -;
 DR EMBL: AF158101; AAD42545.1; -;
 DR EMBL: M22767; AAA88469.1; -;
 DR EMBL: M69268; AAA32544.1; -;
 DR EMBL: Y00148; CAA68343.1; -;
 DR PIR: A00576; XUBP84.
 DR PDB: 1BGF; 30-SEP-94.
 DR PDB: 1BGU; 30-SEP-94.
 DR PDB: 2BGF; 09-DEC-95.

DR PDB: 2BGF; 09-DEC-95.
 DR PDB: 1OKJ; 28-JUL-99.
 DR PDB: 1C3J; 09-AUG-99.
 KW Transferase; Glucosyltransferase; 3D-structure.
 FT CONFLICT 85 KFWA -> NIMO (IN REF. 3).
 SQ SEQUENCE 351 AA; 40666 MW; D1F42B5FE6B9D61 CRC64;
 Query Match 75.0%; Score 30; DB 1; Length 351;
 Best Local Similarity 71.4%; Pred. No. 43;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MOSAFK 7
 DQ 341 MODAFK 347
 RESULT 14
 PAR4_HUMAN STANDARD; PRT; 385 AA.
 AC 096R10; 076067;
 DT 15-JUN-2002 (Ref. 41, Created)
 DT 15-JUN-2002 (Ref. 41, Last sequence update)
 DT 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
 DE like 3) (Coagulation factor II receptor-like 3).
 GN P2RL3 OR PAR4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98389762; PubMed-9722561;
 RA Kahn M.L., Hammes S.R., Botka C., Coughlin S.R.;
 RT "Gene and locus structure and chromosomal localization of the
 RT protease-activated receptor gene family.";
 RL J. Biol. Chem. 273:23290-23296(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98379991; PubMed-9716134;
 RA Kahn M.L., Zheng Y.-W., Huang W., Bigornia V., Zeng D., Moff S.,
 RA Farese R.V., Jr., Tam C., Coughlin S.R.;
 RT "A dual thrombin receptor system for platelet activation.";
 RL Nature 394:690-694(1998).
 RN [3]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-47 AND ARG-68.
 RC TISSUE=lymphoma;
 RX MEDLINE-98283984; PubMed-9618465;
 RA Xu W.-F., Andersen H., Whitmore T.E., Presnell S.R., Yee D.P.,
 RA Ching A., Gilbert T., Davie E.W., Foster D.C.;
 RT "Cloning and characterization of human protease-activated receptor
 RT 4.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:6642-6646(1998).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANTS VAL-296 AND LEU-310.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Peol C.L., Yi Q.,
 RA Nickerson D.A.;
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP FUNCTION.
 RX MEDLINE-99178992; PubMed-10079109;
 RA Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,
 RA Coughlin S.R.;
 RT "Protease-activated receptors 1 and 4 mediate activation of human
 RT platelets by thrombin.";
 RL J. Clin. Invest. 103:879-887(1999).
 CC -1- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
 CC proteins that stimulate phosphoinositide hydrolysis. May play a
 CC role in platelets activation.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in lung,
 CC pancreas, thyroid, testis and small intestine. Not expressed in
 CC brain, kidney, spinal cord and peripheral blood leukocytes. Also

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:44:31 ; Search time 3.05613 Seconds
(Without alignments)
471.946 Million cell updates/sec

Title: US-10-091-442-2

Perfect score: 40

Sequence: 1 MOSAFPK 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	90.0	380	11	Q9D695	09D695 mus musculus
2	90.0	380	11	Q9D695	09D695 mus musculus
3	90.0	380	11	Q9D695	09D695 mus musculus
4	85.0	313	2	Q9D695	09D695 mus musculus
5	85.0	1800	5	Q9D695	09D695 mus musculus
6	82.5	302	10	Q9D695	09D695 mus musculus
7	82.5	397	16	Q9D695	09D695 mus musculus
8	82.5	865	3	Q9D695	09D695 mus musculus
9	80.0	77	2	Q9D695	09D695 mus musculus
10	80.0	161	16	Q9D695	09D695 mus musculus
11	80.0	385	16	Q9D695	09D695 mus musculus
12	80.0	386	3	Q9D695	09D695 mus musculus
13	80.0	386	3	Q9D695	09D695 mus musculus
14	80.0	435	10	Q9D695	09D695 mus musculus
15	80.0	473	5	Q9D695	09D695 mus musculus
16	80.0	712	2	Q9D695	09D695 mus musculus

17	80.0	736	17	Q9D695	09D695 mus musculus
18	80.0	750	16	Q9D695	09D695 mus musculus
19	80.0	1072	5	Q9D695	09D695 mus musculus
20	80.0	1080	5	Q9D695	09D695 mus musculus
21	77.5	112	16	Q9D695	09D695 mus musculus
22	77.5	187	10	Q9D695	09D695 mus musculus
23	77.5	243	16	Q9D695	09D695 mus musculus
24	77.5	275	10	Q9D695	09D695 mus musculus
25	77.5	289	16	Q9D695	09D695 mus musculus
26	77.5	357	16	Q9D695	09D695 mus musculus
27	77.5	357	16	Q9D695	09D695 mus musculus
28	77.5	427	5	Q9D695	09D695 mus musculus
29	77.5	451	5	Q9D695	09D695 mus musculus
30	77.5	514	17	Q9D695	09D695 mus musculus
31	77.5	554	17	Q9D695	09D695 mus musculus
32	77.5	558	17	Q9D695	09D695 mus musculus
33	77.5	570	5	Q9D695	09D695 mus musculus
34	77.5	590	16	Q9D695	09D695 mus musculus
35	77.5	666	16	Q9D695	09D695 mus musculus
36	77.5	729	2	Q9D695	09D695 mus musculus
37	77.5	1108	16	Q9D695	09D695 mus musculus
38	77.5	1145	10	Q9D695	09D695 mus musculus
39	75.0	128	16	Q9D695	09D695 mus musculus
40	75.0	156	2	Q9D695	09D695 mus musculus
41	75.0	163	3	Q9D695	09D695 mus musculus
42	75.0	165	16	Q9D695	09D695 mus musculus
43	75.0	226	8	Q9D695	09D695 mus musculus
44	75.0	228	8	Q9D695	09D695 mus musculus
45	75.0	241	16	Q9D695	09D695 mus musculus

ALIGNMENTS

RESULT 1
ID Q9D695 PRELIMINARY; PRT; 380 AA.
AC Q9D695;
DT 01-JUN-2001 (TREMURel. 17, Created)
DT 01-JUN-2001 (TREMURel. 17, Last sequence update)
DE 4631416M05Rik protein (Megs1n).
GN SERPINB7 OR 4631416M05Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-C57BL/6J; TISSUE=SKIN;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinnagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Katsukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirini L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonard M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmberg L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

RX MEDLINE-21368006; PubMed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangioliproliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
DR EMBL: AK014524; BAB29410.1; -.
DR EMBL: AF105328; AAL16768.1; -.
DR HSSP: P05121; 1A7C.
DR MGD: MGI:2151053; Serpinb7.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin_1.
DR SMART: SM00093; SERPIN_1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 380 AA; 43050 MW; C9240272BCEB9CF4 CRC64;

Query Match          90.0%; Score 36; DB 11; Length 380;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 181 WQSAFTK 187

RESULT 2
ID Q920J5 PRELIMINARY; PRT; 380 AA.
AC Q920J5;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Megsin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21368006; PubMed-11473647;
RA Nangaku M., Miyata T., Suzuki D., Umezono T., Hashimoto T., Wada T.,
RA Yagi M., Nagano N., Inagi R., Kurokawa K.;
RT "Cloning of rodent mesgin revealed its up-regulation in
RT mesangioliproliferative nephritis.";
RL Kidney Int. 60:641-652(2001).
CC 1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AF105329; AAL16769.1; -.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin_1.
DR PROSITE: PS00284; SERPIN; UNKNOWN_1.
KW Serpin.
SQ SEQUENCE 380 AA; 42821 MW; D8076CABEEZC2FBC CRC64;

Query Match          90.0%; Score 36; DB 11; Length 380;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 181 WQSAFTK 187

RESULT 3
ID Q9REE9 PRELIMINARY; PRT; 313 AA.
AC Q9REE9;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 33.2 kDa protein.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.

```

```

OX NCBI_TaxID-375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-110SPC4;
RX MEDLINE-97261868; PubMed-9108282;
RA Minder A.C., Narberhaus F., Babst M., Hennecke H., Fischer H.M.;
RT "The dnaK operon belongs to the sigma32-dependent class of heat shock
RT genes in Bradyrhizobium japonicum.";
RL Mol. Gen. Genet. 254:195-206(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-110SPC4;
RX MEDLINE-20082837; PubMed-10613857;
RA Minder A.C., Fischer H.M., Hennecke H., Narberhaus F.;
RT "Role of HrcA and CIRCE in the heat shock regulatory network of
RT Bradyrhizobium japonicum.";
RL J. Bacteriol. 182:14-22(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-110SPC4;
RX MEDLINE-21150465; PubMed-11251836;
RA Minder A.C., de Rudder K.E.E., Narberhaus F., Fischer H.M.,
RA Hennecke H., Geiger O.;
RT "Phosphatidylcholine levels in Bradyrhizobium japonicum membranes are
RT critical for an efficient symbiosis with the soybean host plant.";
RL Mol. Microbiol. 39:1186-1198(2001).
DR EMBL: Y09633; CAB60667.1; -.
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 33241 MW; 2F530355FCEB1B59 CRC64;

```

```

Query Match          85.0%; Score 34; DB 2; Length 313;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQSAFTK 7
DB 295 WQSAFTK 301

RESULT 4
ID Q94882 PRELIMINARY; PRT; 1800 AA.
AC Q94882; 001381;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-JUN-1998 (TREMblrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE DILUTE CLASS UNCONVENTIONAL MYOSIN.
DE DIDUM OR CG2146.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID-7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Maciver B.;
RL Thesis (1997); University of Edinburgh, Edinburgh, UK.
DR EMBL: Y08160; CAA69352.1; -.
DR HSSP: P10587; 1BR2.
DR FlyBase: FBgn0015933; didum.
DR InterPro: IPR002710; DIL.
DR InterPro: IPR000048; IQ_region.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF01843; DIL_1.
DR Pfam: PF00612; IQ_6.
DR Pfam: PF00063; myosin_head_1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head_1.
DR PRODOM: PD000376; DIL_1.
DR SMART: SM00015; IQ_6.
DR SMART: SM00242; MYSC_1.
DR PROSITE: PS50096; IQ_1.
KW Myosin.

```

SQ SEQUENCE 1800 AA; 208948 MW; E88B0156FAAC57F1 CRC64;

Query Match

Best Local Similarity 71.4%; Score 34; DB 5; Length 1800;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
11:1111
DB 649 WETAFTK 655

RESULT 5

O9LUY1 PRELIMINARY: PRT: 302 AA.

AC O9LUY1:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE C1a-prenyltransferase homolog-like, undecaprenyl phosphate synthetase-like.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty pl and YAC clones.";
RT DNA Res. 7:31-63(2000).
RL EMBL; AB020755; BAA97348.1; -
DR InterPro: IPR001993; Mitochlcarrier.
DR InterPro: IPR001441; UPP_synth.
DR Pfam: PF01255; UPP_synthetase; 1.
DR Prodom: PD003461; UPP_synth; 1.
DR Trifams: TIGR00055; UPPS; 1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE: PS01066; UPP_SYNTHETASE; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 302 AA; 34893 MW; C35020DDDF493ABE CRC64;

Query Match 82.5%; Score 33; DB 10; Length 302;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
11:1111
DB 262 WQSAFTE 268

RESULT 6

O9KFB1 PRELIMINARY: PRT: 397 AA.

AC O9KFB1:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Chromate transporter.
GN CHRA OR BH0574.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takemi H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001509; BAB04293.1; -
DR InterPro: IPR003370; Chromate_transp.
DR Pfam: PF02417; Chromate_transp; 2.
KW Complete proteome.
SQ SEQUENCE 397 AA; 42918 MW; 1E1ICE27DB61ED56 CRC64;

Query Match 82.5%; Score 33; DB 16; Length 397;
Best Local Similarity 71.4%; Pred. No. 92;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
11:1111
DB 160 WQSAFTQ 166

RESULT 7

O43022 PRELIMINARY: PRT: 865 AA.

AC O43022:
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypothetical 99.2 kDa protein C354.08C in chromosome II.
GN SPBC354.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Wood V., Rajandream M.A., Barrell B.G., Hiltbert H., Duesterhoeft A.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: TO A. THALIANA ML23.19.
DR EMBL; AL022071; CAAL7808.1; -
DR InterPro: IPR003864; DUF221.
DR Pfam: PF02714; DUF221; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 32
FT TRANSMEM 107
FT TRANSMEM 152
FT TRANSMEM 172
FT TRANSMEM 375
FT TRANSMEM 395
FT TRANSMEM 512
FT TRANSMEM 532
FT TRANSMEM 583
FT TRANSMEM 603
FT TRANSMEM 641
FT TRANSMEM 661
FT TRANSMEM 666
FT TRANSMEM 686
FT TRANSMEM 9206
SQ SEQUENCE 865 AA; 99206 MW; 9F7BD97BA43FB1904 CRC64;

Query Match 82.5%; Score 33; DB 3; Length 865;
Best Local Similarity 71.4%; Pred. No. 21e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WQSAFTK 7
11:1111
DB 448 WQSAFTK 454.

RESULT 8

O52995 PRELIMINARY: PRT: 77 AA.

AC O52995:
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PCOLD-157 mbdD protein.
GN PCOLD-157 MBDD.

```

OS Escherichia coli
OC Plasmid pCOLD-157
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL40 CURED;
RA MEDLINE=98092211; PubMed=9431914;
RA Hofinger C., Karch H., Schmidt H.;
RT "Structure and function of plasmid pCOLD157 of enterohemorrhagic
RT Escherichia coli O157 and its distribution among strains from patients
RT with diarrhea and hemolytic-uremic syndrome.";
RL J. Clin. Microbiol. 36:24-29(1998).
DR EMBL; Y10412; CAA71431.1; -
KW Plasmid.
SQ SEQUENCE 77 AA; 9239 MW; 998FB8A10AD8D5CB CRC64;

Query Match 80.0%; Score 32; DB 2; Length 77;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
Db 26 WSAFT 31

RESULT 9
O92RF9 PRELIMINARY; PRT; 161 AA.
AC O92RF9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical peroxiredoxin protein SMC00072.
GN R00915 OR SMC00072.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
RA Pohl T., Portetelle P., Puehler A., Purnelle B., Ranspenger U.,
RA Renard C., Thebaud P., Vandebol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591785; CAC45487.1; -
DR InterPro: IPR000866; Ahpc-TSA.
DR Pfam; PF00578; Ahpc-TSA; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 161 AA; 16997 MW; 9FAA1CBFD1D31C01 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 161;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WSAFT 7
Db 102 WSAFT 108

RESULT 10
O98B00 PRELIMINARY; PRT; 369 AA.
AC O98B00;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Amino transferase.
GN MRS786.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFE303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003007; BAB52172.1; -
DR InterPro: IPR004839; Amino transferase/2.
DR InterPro: IPR001917; NHTransf.2.
DR Pfam; PF00155; aminotran_1_2; 1.
DR TIGRfams; TIGR01141; hsc; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
KW transferase; Complete proteome.
SQ SEQUENCE 369 AA; 40710 MW; DCB23E0507165570 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 369;
Best Local Similarity 71.4%; Pred. No. 146+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WSAFT 7
Db 315 WSAFT 321

RESULT 11
O9UQYO PRELIMINARY; PRT; 385 AA.
AC O9UQYO;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE O-methyltransferase.
GN Mt-I.
OS Aspergillus parasiticus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
OX NCBI_TaxID=5067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTAH-26;
RX MEDLINE=20011239; PubMed=10543813;
RA Motomura M., Chihaya N., Shinozawa T., Hamasaki T., Yabe K.;
RT "Cloning and characterization of the O-methyltransferase I gene (dmfA)
RT from Aspergillus parasiticus associated with the conversions of
RT demethylsterigmatocystin to sterigmatocystin and
RT dihydrodemethylsterigmatocystin to dihydrosterigmatocystin in
RT aflatoxin biosynthesis.";
RL Appl. Environ. Microbiol. 65:4987-4994(1999).
DR EMBL; AB022906; BAA86104.1; -
DR EMBL; AB022905; BAA86103.1; -
DR InterPro: IPR00183; Decarboxylase2.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR001077; O_methyltransf.
DR Pfam; PF00891; Methyltransf_2; 1.
DR PROSITE; PS00879; ODR_DC_2; UNKNOWN_1.
KW Methyltransferase; Transferase.
FT INT. MET 0
SQ SEQUENCE 385 AA; 43023 MW; C9866E7A96272005 CRC64;

Query Match 80.0%; Score 32; DB 3; Length 385;

```


Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
11:111

DB 330 WQAFT 335

RESULT 12
O9UV62

ID O9UV62 PRELIMINARY; PRT: 386 AA.

AC O9UV62: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, last annotation update)

DE O-methyltransferase B.

GN OMTB.

OS Aspergillus parasiticus.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Euceliales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_Taxid=5067;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SU-1;

RX MEDLINE=20267858; PubMed=10806361;

RA Yu J., Moloshuk C.P., Bhattacharya D., Cleveland T.E.;

RT "Cloning and characterization of avfa and omtb genes involved in aflatoxin biosynthesis in three Aspergillus species.";

RL Gene 248:157-167(2000).

DR EMBL: AF154050; AAF25603.1; "

DR InterPro: IPR000183; Decarboxylase2.

DR InterPro: IPR001077; O-methyltransf.

DR Pfam: PF00891; Methyltransf_2; 1.

DR PROSITE: PS00879; ODR_DC_2; UNKNOWN_1.

DR Methyltransferase; Transferase.

SO SEQUENCE 386 AA; 43142 MW; 13D6F1ED63CF63D CRC64;

Query Match 80.0%; Score 32; DB 3; Length 386;

Best Local Similarity 83.3%; Pred. No. 1.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
11:111

DB 331 WQAFT 336

RESULT 13
O9P900

ID O9P900 PRELIMINARY; PRT: 386 AA.

AC O9P900: 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, last annotation update)

DE O-methyltransferase B.

GN OMTB.

OS Aspergillus flavus.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Euceliales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_Taxid=5059;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CR01-2B;

RX MEDLINE=20267858; PubMed=10806361;

RA Yu J., Moloshuk C.P., Bhattacharya D., Cleveland T.E.;

RT "Cloning and characterization of avfa and omtb genes involved in aflatoxin biosynthesis in three Aspergillus species.";

RL Gene 248:157-167(2000).

DR EMBL: AF159789; AAF26223.1; "

DR InterPro: IPR000183; Decarboxylase2.

DR InterPro: IPR001077; O-methyltransf.

DR Pfam: PF00891; Methyltransf_2; 1.

DR PROSITE: PS00879; ODR_DC_2; UNKNOWN_1.

RN Methyltransferase; Transferase.

SO SEQUENCE 386 AA; 43160 MW; F199017CC28DE9C1 CRC64;

Query Match 80.0%; Score 32; DB 3; Length 386;

Best Local Similarity 83.3%; Pred. No. 1.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
11:111

DB 331 WQAFT 336

RESULT 14
O9FV0

ID O9FV0 PRELIMINARY; PRT: 435 AA.

AC O9FV0: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)

DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)

DE Similarity to GTPase activating protein.

OS Arabidopsis thaliana (mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OX NCBI_Taxid=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;

RX MEDLINE=97471969; PubMed=9330910;

RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,

RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned

RT P1 clones.";

RL DNA Res. 4:215-230(1997).

DR EMBL: AB005232; BAB08757.1; "

DR InterPro: IPR000195; RabGAP_TBC.

DR Pfam: PF00566; TBC; 1.

DR SMART: SM00164; TBC; 1.

SO SEQUENCE 435 AA; 49768 MW; A967D4DC4F703C07 CRC64;

Query Match 80.0%; Score 32; DB 10; Length 435;

Best Local Similarity 83.3%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WSAFT 6
11:111

DB 40 WQAFT 45

RESULT 15
O9GRW8

ID O9GRW8 PRELIMINARY; PRT: 473 AA.

AC O9GRW8: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)

DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)

DE GpII protein.

GN GpII.

OS Drosophila virilis (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OX NCBI_Taxid=7244;

RN [1]

RP SEQUENCE FROM N.A.

RA Campesano S., Chalmers D., Sandrelli F., Megighian A., Pelxoto A.A.,

RT Costa R., Kyriacou C.P.;

RT "Comparative analysis of the nonA region in Drosophila identifies

RT highly diverged 5' gene that may constrain nonA promoter evolution.";

RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AJ298996; CAC10057.1; "

DR FlyBase; FBgn0042731; Dvir\Gp11.
SQ SEQUENCE 473 AA; 55085 MW; A6E3E375C5AB348A CRC64;

Query Match 80.0%; Score 32; DB 5; Length 473;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WQSAFTRK 7
111 : 11
Db 202 WQSIYTRK 208

Search completed: July 11, 2003, 11:58:51
Job time : 6.05613 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:42:16 ; Search time 10.9023 seconds
(without alignments)
232.223 Million cell updates/sec

Title: US-10-091-442-3

Perfect score: 100

Sequence: 1 VERVDFTHLEDTRRINK 19

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

- 1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	380	15	AA848379 Human megakaryocyte
2	100	100.0	380	20	AAV08254 Human megaslin prote
3	100	100.0	380	21	AA824142 Human megaslin prote
4	100	100.0	380	22	AA83075 Human megaslin prote
5	82	82.0	19	15	AA857109 Human megakaryocyte
6	71	71.0	368	20	AAV08256 Mouse megaslin prote
7	71	71.0	368	21	AA824151 Mouse megaslin prote
8	71	71.0	368	22	AA83077 Murine megaslin prot
9	68	68.0	380	20	AAV08255 Rat megaslin protein
10	68	68.0	380	21	AA824150 Rat megaslin protein

11	68	68.0	380	22	AA84286
12	68	68.0	380	22	AA83076
13	50	50.0	390	13	AA82527
14	50	50.0	390	18	AA815241
15	50	50.0	390	20	AAV32077
16	50	50.0	390	20	AAV32078
17	50	50.0	390	20	AAV25928
18	50	50.0	390	22	AAV72654
19	47	47.0	503	16	AA872601
20	47	47.0	503	16	AA872508
21	47	47.0	504	13	AA82482
22	47	47.0	504	16	AA872507
23	47	47.0	504	16	AA872509
24	47	47.0	504	16	AA872505
25	47	47.0	1047	22	AA872507
26	47	47.0	1047	22	AA872491
27	46	46.0	375	15	AA850938
28	46	46.0	375	17	AA86882
29	46	46.0	375	19	AA869287
30	46	46.0	375	20	AAV02607
31	46	46.0	390	18	AA815242
32	46	46.0	390	20	AAV25927
33	46	46.0	390	22	AAV72655
34	46	46.0	410	22	AA804102
35	46	46.0	427	22	AA867060
36	46	46.0	579	23	AA897393
37	46	46.0	732	17	AA800466
38	46	46.0	732	18	AA836105
39	46	46.0	732	18	AA836106
40	46	46.0	1120	22	ABG28279
41	45	45.0	45	22	AB843663
42	45	45.0	45	22	AB826613
43	45	45.0	45	22	AA864615
44	45	45.0	45	22	AA877400
45	45	45.0	45	22	AA821329

ALIGNMENTS

RESULT 1	AA848379	standard; Protein; 380 AA.
ID	AA848379	
XX	AA848379;	
AC	16-AUG-1994 (first entry)	
XX		
DT		
XX		
DE	Human megakaryocyte differentiation factor.	
XX		
KW	Human megakaryocyte differentiation factor; MDF; thrombopoietin;	
KW	haematopoietic stimulating factor; thrombocytopoietin; platelet;	
XX	bone marrow transplantation; cancer chemotherapy.	
OS	Homo sapiens.	
XX		
PN	EP583884-A.	
XX		
PD	23-FEB-1994.	
XX		
PF	19-JUL-1993; 93EP-0305654.	
XX		
PR	17-JUL-1992; 92JP-0212305.	
XX		
PR	04-MAR-1993; 93JP-0067339.	
XX		
PA	(SUNR) SUNTORY LTD.	
XX	(TSUJ) TSUJIMOTO M.	
PI	Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;	
XX	TSUJimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;	
DR	WPI: 1994-058782/08.	
DR	N-PSDB: AA056670.	

Rat megaslin protein
Rat megaslin protein
SCC antigen. Synt
Peptidase type
Hepatitis B virus
Hepatitis B virus
Human SCCA1 protei
Human squamous cel
Truncated human ch
Rat cholesterol 7
Cholesterol 7-alpha
Human cholesterol
Hamster cholesterol
Human cholesterol
Human protein SRO
Human protein SRO
Maspin. Homo sapi
Maspin. Homo sapi
Tumour suppressor
Maspin. Homo sapi
Human maspin prote
Peptidase type
Human SCCA2 protei
Human squamous cel
Botulinum toxin hea
Drosophila melanog
Novel human protei
Bovine Factor-XIII
Wild type bovine f
Mutant bovine fact
Novel human diagno
Peptide #1169 enc
Protein #8612 enc
Human brain expres
Human bone marrow
Peptide #763 enco

```

XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
PS Claim 7; Page 30-32; 47pp; English.
XX
CC Human MDF can be isolated from a culture of human epidermoid
CC carcinoma A431 cells in protein-free medium. The MDF stimulates
CC differentiation of megakaryocytes from myeloid cells in the presence
CC of IL-3. The MDF acts in vivo as a thrombopoietin making it useful
CC for treatment of diseases involving a decrease in platelet number
CC (esp. thrombocytopenia) such as occurs in bone marrow
CC transplantation and in chemotherapy. MDF has mol.wt. 55-57kd by
CC SDS-PAGE and exhibits an isoelectric point of 6.0-7.0.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 100; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLEDTRRNINK 19
DB 126 VERVDFTNHLEDTRRNINK 144

RESULT 2
AAV08254
ID AAV08254 standard; Protein: 380 AA.
XX
AC AAV08254;
XX
DT 14-JUL-1999 (first entry)
XX
DE Human megsin protein.
XX
KW Megsin; mesangial cell; treatment; diagnosis; disease; IGA nephropathy;
KW human; rat; murine.
XX
OS Homo sapiens.
XX
PN W09915652-A1.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-JP04269.
XX
PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI; 1999-276983/23.
DR N-PSDB; AAX56712.
XX
PT Megsin protein expressed specifically in mesangial cells
XX
PS Claim 1; Page 62-64; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megsin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as IGA nephropathy.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 100; DB 20; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 VERVDFTNHLEDTRRNINK 19
DB 126 VERVDFTNHLEDTRRNINK 144

RESULT 3
AAB24142
ID AAB24142 standard; Protein: 380 AA.
XX
AC AAB24142;
XX
DT 30-JAN-2001 (first entry)
XX
DE Human megsin protein sequence SEQ ID NO:2.
XX
KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
KW IGA; immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
OS Homo sapiens.
XX
PN W0200057189-A1.
XX
PD 28-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-JP01646.
XX
PR 19-MAR-1999; 99JP-0075305.
PR 28-OCT-1999; 99JP-0306623.
XX
PA (KURO/) KUROKAWA K.
PA (FUSO ) FUSO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI; 2000-611642/58.
DR N-PSDB; AAB9294.
XX
PT Evaluating renal function comprises assaying megsin protein in
PT biological sample
XX
PS Example 2; Page 66-69; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megsin protein in biological sample. Also
CC described are: (1) use of a anti-megsin protein antibody for diagnosing
CC renal function; and (2) a kit for detecting megsin protein comprising:
CC (a) anti-megsin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying megsin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence represents the human megsin
CC protein, which is given in the exemplification of the present invention.
XX
SQ Sequence 380 AA;
Query Match 100.0%; Score 100; DB 21; Length 380;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTNHLEDTRRNINK 19
DB 126 VERVDFTNHLEDTRRNINK 144

RESULT 4
AAB83075
ID AAB83075 standard; Protein: 380 AA.
XX
AC AAB83075;

```

```

XX 10-JUL-2001 (first entry)
DT Human megasin protein.
DE
XX
XX Human; megasin; mesangial cell proliferative nephritis; nephrotropic;
KW transgenic mouse; glomerular disease; animal model; drug screening.
XX
XX Homo sapiens.
OS
XX WO200124628-A1.
PN
XX
XX 12-APR-2001.
PD
XX
XX 06-OCT-2000; 2000WO-JP06988.
PF
XX
XX 06-OCT-1999; 99JP-0285736.
PR
XX
XX (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
XX Miyata T;
PI
XX WPI: 2001-300136/31.
DR N-PSDB: AAF82438.
XX
XX Mouse model for mesangial cell proliferative nephritis for development
PT and screening of new treatments.
XX
XX Example 4; Page 44-46; 62pp; Japanese.
XX
XX The present sequence is human megasin. The human megasin coding
CC sequence may be introduced into a mouse to produce an animal model of
CC mesangial cell proliferative nephritis. The symptoms include
CC enlargement of the mesangial base region, sedimentation of an immune
CC complex and an increase in mesangial cells. The animal model is useful
CC for analysing the pathology of chronic glomerular diseases and for
CC screening compositions for prevention and treatment of the diseases.
CC Highly uniform models can be made easily and in large numbers using
CC this method.
XX
XX Sequence 380 AA:
SQ
Query Match 100.0%; Score 100; DB 22; Length 380;
Best Local Similarity 100.0%; Pred. No. 7,66-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VERVDFTNHLNLEDTNRINK 19
DB 126 VERVDFTNHLNLEDTNRINK 144
RESULT 5
AAR57109
ID AAR57109 standard; peptide; 19 AA.
XX
XX AAR57109;
AC
XX
XX 16-AUG-1994 (first entry)
DT Human megakaryocyte differentiation factor peptide 3.
XX
XX Human megakaryocyte differentiation factor; MDP; thrombopoietin;
KW haematopoietic stimulating factor; thrombocytopenia; platelet;
XX bone marrow transplantation; cancer chemotherapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 3 /note= "not determined"
FT Misc-difference 14 /note= "not determined"
FT

```

```

FT Misc-difference 15 /note= "not determined"
FH
XX
XX EP583884-A.
PN
XX
XX 23-FEB-1994.
PD
XX
XX 19-JUL-1993; 93EP-0305654.
PF
XX
XX 17-JUL-1992; 92JP-0212305.
PR 04-MAR-1993; 93JP-0067339.
XX
XX (SUNR ) SUNTORY LTD.
PA (TSUJ/) TSUJIMOTO M.
XX
XX Ishida N, Iwasa F, Kurihara T, Miura K, Nakazato H;
PI Tsujimoto M, Tsuruoka N, Yamaguchi N, Yamachi K;
XX
XX WPI: 1994-058782/08.
DR
XX
XX New megakaryocyte differentiation factor - isolated from human
PT epidermoid carcinoma cells, used to treat conditions involving a
PT decrease in platelets
XX
XX Claim 1; Page 18; 47pp; English.
XX
XX Human MDF (see AAR48379) can be isolated from a culture of human
CC epidermoid carcinoma A431 cells in protein-free medium. The MDF
CC stimulates differentiation of megakaryocytes from myeloid cells
CC in the presence of IL-3. The MDF acts in vivo as a thrombopoietin
CC making it useful for treatment of diseases involving a decrease
CC in platelet number (esp. thrombocytopenia) such as occurs in bone
CC marrow transplantation and in chemotherapy. MDF has mol.wt.
CC 55-57kD by SDS-PAGE, exhibits an isoelectric point of 6.0-7.0 and
CC contains an amino acid sequence comprising at least one of the
CC sequences AAR57107-R57115.
XX
XX Sequence 19 AA:
SQ
Query Match 82.0%; Score 82; DB 15; Length 19;
Best Local Similarity 84.2%; Pred. No. 1,86-06;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 VERVDFTNHLNLEDTNRINK 19
DB 1 VEXVDFTNHLNLEDTXNINK 19
RESULT 6
AAV08256
ID AAV08256 standard; Protein; 368 AA.
XX
XX AAV08256;
AC
XX
XX 14-JUL-1999 (first entry)
DT Mouse megasin protein.
XX
XX Mouse megasin protein.
KW Megasin; mesangial cell; treatment; diagnosis; disease; IgA nephropathy;
XX human; rat; murine.
XX
XX Mus musculus.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..368
FT /note= "partial sequence"
FT
XX
XX WO9915652-A1.
PN
XX
XX 01-APR-1999.
PD
XX
XX 22-SEP-1998; 98WO-JP04269.
PF
XX

```

```

PR 22-SEP-1997; 97JP-0275302.
XX
PA (KURO/) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 1999-276983/23.
XX N-PSDB: AAX56714.
XX
PI Megsin protein expressed specifically in mesangial cells
XX
PS Claim 1; Page 76-79; 100pp; Japanese.
XX
CC This invention describes the isolation of novel megin nucleic acid and
CC proteins from human, rat and mouse tissue. This protein is expressed
CC specifically in mesangial cells. The products of the invention are
CC useful for the treatment and diagnosis of diseases involving mesangial
CC cells, such as Iga nephropathy.
XX
SQ Sequence 368 AA:
XX
Query Match 71.0%; Score 71; DB 20; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0034;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0
OY 1 VERVDFTHLEPTRRINK 19
| | | | | | | : | | | | |
DB 114 VERVDFTHNDVPTREKINK 132

RESULT 7
AAB24151
ID AAB24151 standard; Protein; 368 AA.
XX
AC AAB24151;
XX
DT 30-JAN-2001 (first entry)
XX
DE Mouse megin protein sequence SEQ ID NO:21.
XX
KW Megsin; mesangium-predominant gene; serpin regulated; nephropathy;
KW Iga; immunoglobulin A; detection; renal function; renal disorder;
KW diagnosis; biological sample; blood; urine.
XX
OS Mus musculus.
XX
PN WO200057189-A1.
XX
PD 28-SEP-2000.
XX
PF 17-MAR-2000; 2000WO-JP01646.
XX
PR 19-MAR-1999; 99JP-0075305.
XX
PR 28-OCT-1999; 99JP-0306623.
XX
PA (KURO/) KUROKAWA K.
PA (FUJO ) FUJO PHARM IND LTD.
PA (MIYA/) MIYATA T.
XX
PI Miyata T;
XX
DR WPI: 2000-611642/58.
DR N-PSDB: AAC55239.
XX
PT Evaluating renal function comprises assaying megin protein in
PT biological sample
XX
PS Disclosure: Page 89-91; 93pp; Japanese.
XX
CC The present invention describes a method for evaluating renal function.
CC The method comprises assaying megin protein in biological sample. Also
CC described are: (1) use of a anti-megsin protein antibody for diagnosing

```

```

CC renal function; and (2) a kit for detecting mesgin protein compising:
CC (a) anti-mesgin protein antibody attached to solid magnetic particles;
CC (b) direct or indirect fixing for the antibody to the particles; and
CC (c) a magnet. The process is useful for evaluating renal function and
CC diagnosing renal disorders by assaying mesgin protein in biological
CC samples (preferably urine or blood). The process is reproducible and
CC gives accurate results. The present sequence represents the mouse mesgin
CC protein, which is given in the exemplification of the present invention.
CX
XX
SO Sequence 368 AA:

Query Match 71.0%; Score 71; DB 21; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0034;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0

QY 1 VERVDFTNHLEDTRRNINK 19
   ||||| :||| ||
DB 114 VERVDFTNDVQDTRFKINK 132

RESULT 8
AAB83077
ID AAB83077 standard; Protein; 368 AA.
AC AAB83077;
XX 10-JUL-2001 (first entry)
DT
DE Murine mesgin protein.
XX
XX Mouse; mesgin; mesangial cell proliferative nephritis; nephrotropic;
XX transgenic mouse; glomerular disease; animal model; drug screening.
XX
XX Mus musculus.
OS
XX WO200124628-A1.
PN
XX 12-APR-2001.
PD
XX 06-OCT-2000; 2000WO-JP06988.
PF
XX 06-OCT-1999; 99JP-0285736.
PR
XX (KUROO) KUROKAWA K.
PA (MIYA/) MIYATA T.
XX
XX Miyata T;
PI
XX
XX WPI: 2001-300136/31.
DR N-PSDB; AAF82440.
XX
XX Mouse model for mesangial cell proliferative nephritis for development
XX and screening of new treatments
XX
XX Disclosure: Page 52-53; 62pp; Japanese.
PS
XX
XX The present sequence is murine mesgin protein. The human mesgin coding
XX sequence may be introduced into a mouse to produce an animal model of
XX mesangial cell proliferative nephritis. The symptoms include
XX enlargement of the mesangial base region, sedimentation of an immune
XX complex and an increase in mesangial cells. The animal model is useful
XX for analysing the pathology of chronic glomerular diseases and for
XX screening compositions for prevention and treatment of the diseases.
XX Highly uniform models can be made easily and in large numbers using
XX this method.
XX
XX Sequence 368 AA:

Query Match 71.0%; Score 71; DB 22; Length 368;
Best Local Similarity 73.7%; Pred. No. 0.0034;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 VERVDFTNHLEDTRRNINK 19

```

Db 114 VERVDFTNVDTRFKINK 132

|||||||::||| |||

RESULT 9

AAV08255

ID AAV08255 standard; Protein: 380 AA.

XX

AC AAV08255;

XX

DT 14-JUL-1999 (first entry)

XX

DE Rat megalin protein.

XX

KW Megalin; mesangial cell; treatment; diagnosis; disease; Iga nephropathy;

KV human; rat; murine.

XX

OS Rattus rattus.

XX

PN WO915652-A1.

XX

PD 01-APR-1999.

XX

PE 22-SEP-1998; 98WO-JP04269.

XX

PR 22-SEP-1997; 97JP-0275302.

XX

PA (KURO/) KUROKAWA K.

PA (MIYA/) MIYATA T.

XX

PI Miyata T;

XX

DR WPI: 1999-276983/23.

DR N-PSDB: AAX56712.

XX

PT Megalin protein expressed specifically in mesangial cells

PS

PS Claim 1; Page 69-72; 100pp; Japanese.

CC

CC This invention describes the isolation of novel megalin nucleic acid and

CC proteins from human, rat and mouse tissue. This protein is expressed

CC specifically in mesangial cells. The products of the invention are

CC useful for the treatment and diagnosis of diseases involving mesangial

CC cells, such as Iga nephropathy.

XX

SO Sequence 380 AA;

QY

Query Match 68.0%; Score 68; DB 20; Length 380;

Best Local Similarity 68.4%; Pred. No. 0.011;

Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 126 VERVDFTNVDTRFKINK 144

1 VERVDFTNHLEDTRRNINK 19

|||||||::||| |||

AAAB24150

ID AAB24150 standard; Protein: 380 AA.

XX

AC AAB24150;

XX

DT 30-JAN-2001 (first entry)

XX

DE Rat megalin protein sequence SEQ ID NO:19.

XX

KW Megalin; mesangium-predominant gene; serpin regulated; nephropathy;

KV Iga; immunoglobulin A; detection; renal function; renal disorder;

XX

OS Rattus norvegicus.

XX

XX

Key Location/Qualifiers

FT Misc-difference 51

FT /note= "unspecified"

FT Misc-difference 94

FT /note= "unspecified"

XX

PN WO200057189-A1.

XX

PD 28-SEP-2000.

XX

PE 17-MAR-2000; 2000WO-JP01646.

XX

PR 19-MAR-1999; 99JP-0075305.

PR 28-OCT-1999; 99JP-0306623.

XX

PA (KURO/) KUROKAWA K.

PA (FUSO) FUSO PHARM IND LTD.

PA (MIYA/) MIYATA T.

XX

PI Miyata T;

XX

DR WPI: 2000-611642/58.

DR N-PSDB: AAC35238.

XX

PT Evaluating renal function comprises assaying megalin protein in

PT biological sample -

XX

PS Example 2; Page 81-84; 93pp; Japanese.

XX

CC The present invention describes a method for evaluating renal function.

CC The method comprises assaying megalin protein in biological sample. Also

CC described are: (1) use of an anti-megalin protein antibody for diagnosing

CC renal function; and (2) a kit for detecting megalin protein comprising:

CC (a) anti-megalin protein antibody attached to solid magnetic particles;

CC (b) direct or indirect fixing for the antibody to the particles; and

CC (c) a magnet. The process is useful for evaluating renal function and

CC diagnosing renal disorders by assaying megalin protein in biological

CC samples (preferably urine or blood). The process is reproducible and

CC gives accurate results. The present sequence represents the rat megalin

CC protein, which is given in the exemplification of the present invention.

XX

SO Sequence 380 AA;

QY

Query Match 68.0%; Score 68; DB 21; Length 380;

Best Local Similarity 68.4%; Pred. No. 0.011;

Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 126 VERVDFTNVDTRFKINK 144

1 VERVDFTNHLEDTRRNINK 19

|||||||::||| |||

AAAG64286

ID AAG64286 standard; Protein: 380 AA.

XX

AC AAG64286;

XX

DT 21-SEP-2001 (first entry)

XX

DE Rat megalin protein.

XX

KW Rat; megalin; renal mesangial cell; mesangium proliferative nephritis.

KV Rattus norvegicus.

XX

PN WO200148019-A1.

XX

PD 05-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-JP09251.

XX

PR 28-DEC-1999; 99JP-0373677.

XX

PT New DNA encoding squamous cell carcinoma antigen - allowing large
 PT scale antigen prodn. for use as diagnostic reagent or immunogen,
 PT also useful as hybridisation probe

PS Claim 2: Page 9-10, Fig 4; 21pp; German.

XX The protein relates to an antigen (Ag) associated with squamous cell
 CC carcinoma (SCC). The Ag (which can be isolated from SCC of the
 CC cervix uteri, pulmonary or oesophageal carcinoma, etc.) is known as a
 CC diagnostic reagent for SCC but is not easily produced on a large scale
 CC from human cancer cells. By expressing the DNA sequence encoding this
 CC protein, in e.g. E.coli, yeast or mammalian cells, Ag can be prep'd. In
 CC large quantities. Ag can also be used as an immunogen (to produce
 CC monoclonal antibodies, also useful diagnostically) while the Ag-
 CC encoding gene can be detected by hybridisation with this sequence.
 CC Ag is homologous with serine protease inhibitors so may be useful
 CC therapeutically.
 SQ Sequence 390 AA;

Query Match 50.0%; Score 50; DB 13; Length 390;
 Best Local Similarity 55.6%; Pred. No. 8.6;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNIN 18
 ||||| |::|:
 DB 131 VESVDFANAPESRKIN 148

RESULT 14

AAW15241
 ID AAW15241 standard; Protein: 390 AA.

AC AAW15241;

DT 21-JUL-1997 (first entry)

DE Psoriasis type I.

XX Psoriasis; psoriasis; cancer; cell proliferation; therapy;
 KW antagonist.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Active-site 357

FT /note= "Thr-357 may correspond to the P'2 or P'3
 FT position of reactive site of homologous
 FT serpins"

PN MO9714425-A1.

PD 24-APR-1997.

XX 16-OCT-1996; 96WO-US16599.

XX 20-MAR-1996; 96US-0013755.

PR 17-OCT-1995; 95US-0005679.

XX (GEHO) GEN HOSPITAL CORP.

PA (SHIS) SHISEIDO CO LTD.

DR Baeu PC, Goetlinck PF, Hibino T, Takahashi T;

XX WPI: 1997-244853/22.

DR N-PSDB: AAT65948.

PT Modulating cell proliferation or apoptosis - by modulating activity
 PT of psoriasis type I and II genes, e.g. using (ant)agonists,
 PT useful for treatment of cancer or psoriasis

PS Disclosure: Page 44-45; 63pp; English.

XX Psoriasis type I (AAW15241) is a 43 kDa polypeptide that is

CC expressed in psoriatic tissue, e.g. psoriatic epidermis, at least 2
 CC and pref. at least 5-10 times more abundantly than in normal
 CC tissue. It is a cross-class inhibitor, inhibits cathepsin L less
 CC efficiently than does squamous cell carcinoma-antigen (SCC-A), does
 CC not inhibit cathepsin B or cathepsin H, is active at pH 5.0, and is
 CC secreted. Its amino acid sequence, deduced from an isolated cDNA
 CC clone (AAT65948), shows 98% homology to SCC-A. A related sequence,
 CC psoriasis type II (AAW15242), has also been identified. These
 CC polypeptides can be used to identify psoriasis antagonists useful
 CC in the treatment of cell proliferation disorders, e.g. cancer and
 CC psoriasis.

SQ Sequence 390 AA;
 Query Match 50.0%; Score 50; DB 18; Length 390;
 Best Local Similarity 55.6%; Pred. No. 8.6;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 VERVDFTNHLDETRRNIN 18
 ||||| |::|:
 DB 131 VESVDFANAPESRKIN 148

RESULT 15

AAV32077
 ID AAV32077 standard; Protein: 390 AA.

AC AAV32077;

DT 17-JAN-2000 (first entry)

DE Hepatitis B virus receptor SCCL.

XX SCCL; squamous cell carcinoma antigen; human; hepatitis B virus;
 KW HBV; receptor; antiviral; virucide; screening; therapy.

OS Homo sapiens.

PN EP952217-A2.

PD 27-OCT-1999.

XX 03-MAR-1999; 99EP-0200603.

PR 12-MAR-1998; 98IT-MI00498.

XX (ITUY-) ITAL MIN UNIV RICERCA SCI & TECNOLOGICA.

PA Fassina G, De Falco S, Verdoliva A, Ruvo M;

DR WPI: 1999-602802/52.

DR N-PSDB: AAZ20257.

PT Use of a recombinant protein and its allelic variant as receptors for a
 PT hepatitis virus, useful for producing and screening for antiviral drugs

PS Claim 1: Fig 1; 13pp; English.

XX The present sequence comprises human squamous cell carcinoma
 CC antigen (SCCL), a protein newly identified as being a receptor for
 CC hepatitis B virus (HBV). Recombinant proteins having fully, or
 CC in part, the present sequence or an allelic variant (see AAV32078)
 CC of it, are claimed and can be produced in claimed prokaryotic
 CC or eukaryotic (especially human HepG2 cell line) host cells
 CC using claimed vectors carrying SCCL nucleic acids (see AAZ20257 and
 CC AAZ20258). Transgenic animals expressing SCCL or its allelic
 CC variants are also claimed. The recombinant SCCL proteins are used
 CC in molecular screening assays to identify new drugs, as drugs for
 CC the treatment of hepatitis B and in the production of monoclonal
 CC and polyclonal antibodies (claimed). The transgenic animals allow
 CC the possibility of in vivo evaluation of the effect of compounds
 CC with anti-hepatitis activity.

XX Sequence 390 AA;

Query Match 50.0%; Score 50; DB 20; Length 390;
 Best Local Similarity 55.6%; Pred. No. 8.6;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VERVDFTNHLEDTRRNIN 18
 || || | | : | : ||
 Db 131 VESVDFAFANAPESRKRIN 148

Search completed: July 11, 2003, 11:54:19
 Job time : 11.9023 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:49:36 ; Search time 3.39709 Seconds
(Without alignments)
164.563 Million cell updates/sec

Title: US-10-091-442-3

Perfect score: 100

Sequence: 1 VERVDFTNHLEDRNRINK 19

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_patents-AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
7: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*
8: /cgn2_6/ptodata/1/1aa/6E.COMB.pep:*
9: /cgn2_6/ptodata/1/1aa/6F.COMB.pep:*
10: /cgn2_6/ptodata/1/1aa/6G.COMB.pep:*
11: /cgn2_6/ptodata/1/1aa/6H.COMB.pep:*
12: /cgn2_6/ptodata/1/1aa/6I.COMB.pep:*
13: /cgn2_6/ptodata/1/1aa/6J.COMB.pep:*
14: /cgn2_6/ptodata/1/1aa/6K.COMB.pep:*
15: /cgn2_6/ptodata/1/1aa/6L.COMB.pep:*
16: /cgn2_6/ptodata/1/1aa/6M.COMB.pep:*
17: /cgn2_6/ptodata/1/1aa/6N.COMB.pep:*
18: /cgn2_6/ptodata/1/1aa/6O.COMB.pep:*
19: /cgn2_6/ptodata/1/1aa/6P.COMB.pep:*
20: /cgn2_6/ptodata/1/1aa/6Q.COMB.pep:*
21: /cgn2_6/ptodata/1/1aa/6R.COMB.pep:*
22: /cgn2_6/ptodata/1/1aa/6S.COMB.pep:*
23: /cgn2_6/ptodata/1/1aa/6T.COMB.pep:*
24: /cgn2_6/ptodata/1/1aa/6U.COMB.pep:*
25: /cgn2_6/ptodata/1/1aa/6V.COMB.pep:*
26: /cgn2_6/ptodata/1/1aa/6W.COMB.pep:*
27: /cgn2_6/ptodata/1/1aa/6X.COMB.pep:*
28: /cgn2_6/ptodata/1/1aa/6Y.COMB.pep:*
29: /cgn2_6/ptodata/1/1aa/6Z.COMB.pep:*
30: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
31: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
32: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
33: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*
34: /cgn2_6/ptodata/1/1aa/6E.COMB.pep:*
35: /cgn2_6/ptodata/1/1aa/6F.COMB.pep:*
36: /cgn2_6/ptodata/1/1aa/6G.COMB.pep:*
37: /cgn2_6/ptodata/1/1aa/6H.COMB.pep:*
38: /cgn2_6/ptodata/1/1aa/6I.COMB.pep:*
39: /cgn2_6/ptodata/1/1aa/6J.COMB.pep:*
40: /cgn2_6/ptodata/1/1aa/6K.COMB.pep:*
41: /cgn2_6/ptodata/1/1aa/6L.COMB.pep:*
42: /cgn2_6/ptodata/1/1aa/6M.COMB.pep:*
43: /cgn2_6/ptodata/1/1aa/6N.COMB.pep:*
44: /cgn2_6/ptodata/1/1aa/6O.COMB.pep:*
45: /cgn2_6/ptodata/1/1aa/6P.COMB.pep:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	19	US-08-611-977-3	Sequence 3, Appl1
2	100	100.0	380	US-08-472-659-34	Sequence 34, Appl1
3	100	100.0	380	US-08-474-661-34	Sequence 34, Appl1
4	100	100.0	380	US-08-611-977-34	Sequence 34, Appl1
5	82	82.0	19	US-08-472-659-3	Sequence 3, Appl1
6	82	82.0	19	US-08-474-661-3	Sequence 3, Appl1
7	50	50.0	390	US-08-568-147B-2	Sequence 5, Appl1
8	47	47.0	482	US-08-135-510-5	Sequence 2, Appl1
9	47	47.0	482	US-08-483-852-5	Sequence 5, Appl1
10	47	47.0	482	US-08-477-953-5	Sequence 5, Appl1
11	47	47.0	482	US-08-477-952-5	Sequence 5, Appl1
12	47	47.0	503	US-08-135-511-29	Sequence 29, Appl1
13	47	47.0	503	US-08-187-453-29	Sequence 29, Appl1
14	47	47.0	504	US-08-135-511-28	Sequence 28, Appl1
15	47	47.0	504	US-08-135-511-30	Sequence 30, Appl1
16	47	47.0	504	US-08-483-852-8	Sequence 8, Appl1
17	47	47.0	504	US-08-361-458-3	Sequence 8, Appl1
18	47	47.0	504	US-08-477-953-8	Sequence 8, Appl1
19	47	47.0	504	US-08-187-453-28	Sequence 28, Appl1
20	47	47.0	504	US-08-187-453-30	Sequence 30, Appl1
21	47	47.0	504	US-08-477-952-8	Sequence 8, Appl1
22	47	47.0	506	US-08-845-161A-18	Sequence 18, Appl1
23	47	47.0	506	US-08-270-751-18	Sequence 18, Appl1
24	46	46.0	375	US-08-121-714-2	Sequence 2, Appl1
25	46	46.0	375	US-08-477-108A-2	Sequence 2, Appl1
26	46	46.0	375	US-08-477-112-2	Sequence 2, Appl1
27	46	46.0	375	PCT-US93-08322-2	Sequence 2, Appl1

28	46	46.0	732	5	PCT-US95-17026-2	Sequence 2, Appl1
29	45	45.0	391	4	US-09-123-912-110	Sequence 110, App
30	45	45.0	391	4	US-09-643-597-110	Sequence 110, App
31	45	45.0	400	4	US-09-123-912-112	Sequence 112, App
32	45	45.0	400	4	US-09-643-597-112	Sequence 112, App
33	43	43.0	390	4	US-09-266-910-3	Sequence 3, Appl1
34	43	43.0	390	4	US-09-266-910-4	Sequence 4, Appl1
35	43	43.0	443	4	US-09-134-001C-4915	Sequence 4915, Ap
36	42	42.0	261	4	US-09-110-959A-2	Sequence 2, Appl1
37	42	42.0	317	4	US-09-134-001C-4537	Sequence 4537, Ap
38	41	41.0	379	4	US-09-200-965-2	Sequence 2, Appl1
39	41	41.0	379	1	US-08-121-714-4	Sequence 4, Appl1
40	41	41.0	379	1	US-08-477-108A-4	Sequence 4, Appl1
41	41	41.0	379	2	US-08-477-112-4	Sequence 4, Appl1
42	41	41.0	379	5	PCT-US93-08322-4	Sequence 4, Appl1
43	41	41.0	501	4	US-09-134-001C-5280	Sequence 5280, Ap
44	40.5	40.5	616	4	US-08-965-902A-2	Sequence 2, Appl1
45	40	40.0	279	3	US-08-872-979-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-611-977-3
Sequence 3, Application US/08611977
Patent No. 5972886
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURUOKA, Hiroshi
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5972886uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5972886oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,977
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-611-977-3

Query Match 100.0%; Score 100; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDETRRNK 19
DB 1 VERVDFTNHLDETRRNK 19

RESULT 2

US-08-472-659-34
Sequence 34, Application US/08472659
Patent No. 5831030
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, No. 5831030uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-659-34

Query Match 100.0%; Score 100; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDETRRNK 19
DB 126 VERVDFTNHLDETRRNK 144

RESULT 3

US-08-474-661-34
Sequence 34, Application US/08474661
Patent No. 5874253
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IMASA, Fuyuki
APPLICANT: TSURUOKA, No. 5874253uo
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5874253uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. 5874253omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,661
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: REA, TERESA STANER
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 001560-204
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-661-34

Query Match 100.0%; Score 100; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDETRRNK 19
DB 126 VERVDFTNHLDETRRNK 144

RESULT 4

US-08-611-977-34
; Sequence 34, Application US/08611977
; Patent No. 5972886
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, No. 5972886uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5972886uhltro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5972886oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,977
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/091,028
; FILING DATE: 14-JUL-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 4-212305
; FILING DATE: 17-JUL-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 6-067339
; FILING DATE: 04-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 001560-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-611-977-34

Query Match 100.0%; Score 100; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDTRRNINK 19
DB 126 VERVDFTNHLDTRRNINK 144

RESULT 5
US-08-472-659-3
; Sequence 3, Application US/08472659
; Patent No. 5831030
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, No. 5831030uo
; APPLICANT: NAKAZATO, Hiroshi

APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. 5831030uhltro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
APPLICANT: YAMAGUCHI, No. 5831030oml
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,659
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 5-067339
FILING DATE: 04-MAR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-472-659-3

Query Match 82.0%; Score 82; DB 2; Length 19;
Best Local Similarity 84.2%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VERVDFTNHLDTRRNINK 19
DB 1 VEXVDFTNHLDTJXXNINK 19

RESULT 6
US-08-474-661-3
; Sequence 3, Application US/08474661
; Patent No. 5874253
; GENERAL INFORMATION:
; APPLICANT: TSUJIMOTO, Masafumi
; APPLICANT: IWASA, Fuyuki
; APPLICANT: TSUROOKA, No. 5874253uo
; APPLICANT: NAKAZATO, Hiroshi
; APPLICANT: MIURA, Kenju
; APPLICANT: ISHIDA, No. 5874253uhltro
; APPLICANT: KURIHARA, Tatsuya
; APPLICANT: YAMAGUCHI, Kozo
; APPLICANT: YAMAGUCHI, No. 5874253oml
; TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
; NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Burns, Doane, Swecker & Mathis
;; STREET: George Mason Bldg., Washington & Prince Sts.
;; CITY: Alexandria
;; STATE: Virginia
;; COUNTRY: United States
;; ZIP: 22313-1404
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,661
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;;
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/091,028
;; FILING DATE: 14-JUL-1993
;; APPLICATION NUMBER: JP 4-212305
;; FILING DATE: 17-JUL-1992
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: JP 5-067339
;; FILING DATE: 04-MAR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REA, TERESA STANER
;; REGISTRATION NUMBER: 30,427
;; REFERENCE/DOCKET NUMBER: 001560-204
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 836-6620
;; TELEFAX: (703) 836-6620
;; INFORMATION FOR SEQ ID NO: 3:
;;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 19 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;;
;; MOLECULE TYPE: peptide
;;
;; US-08-474-661-3

Query Match 82.0%; Score 82; DB 2; Length 19;
Best Local Similarity 84.2%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VERVDFTHLEDTNRNINK 19
|| ||||| |||||
Db 1 VEXYDFTHLEDTXNINK 19

RESULT 7
US-08-568-147B-2
; Sequence 2, Application US/08568147B
; Patent No. 5783422
; GENERAL INFORMATION:
; APPLICANT: Sumitani, Yoshinori
; APPLICANT: Kato, Hiroshi
; APPLICANT: Sekiguchi, Kiyoshi
; APPLICANT: Takeda, Katsunichi
; TITLE OF INVENTION: DNA FRAGMENT CODING FOR SQUAMOUS CELL
; TITLE OF INVENTION: CARCINOMA-ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/568,147B
;; FILING DATE:
;; CLASSIFICATION: 536
;;
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: 800,952
;; FILING DATE: 02-DEC-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Digilio, Frank S.
;; REGISTRATION NUMBER: 31,346
;; REFERENCE/DOCKET NUMBER: 8425
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 516-742-4343
;; TELEFAX: 516-742-4366
;; TELE: 230 901 SANS UR
;; INFORMATION FOR SEQ ID NO: 2:
;;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 390 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;;
;; MOLECULE TYPE: protein
;;
;; US-08-568-147B-2

Query Match 50.0%; Score 50; DB 1; Length 390;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VERVDFTHLEDTNRNIN 18
|| ||||| |||||
Db 131 VESYDFANAPESRRKTN 148

RESULT 8
US-08-135-510-5
; Sequence 5, Application US/08135510
; Patent No. 5420028
; GENERAL INFORMATION:
; APPLICANT: CHIANG, John Young Ling
; TITLE OF INVENTION: Truncated Human Cholesterol
; TITLE OF INVENTION: 7a-Hydroxylase, Method of Production and Use Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/135,510
; FILING DATE: 13-OCT-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: SANDERCOCK, Collin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 18748/176 HOCE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELE: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 482 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-135-510-5

Query Match 47.0%; Score 47; DB 1; Length 482;
Best Local Similarity 66.7%; Pred. No. 5.2;

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,952
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/361,458
FILING DATE: 21-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/135,511
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/135,488
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/135,510
FILING DATE: 13-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: SANDERCOCK, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 18748/221 HOCCE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 482 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-477-952-5

Query Match 47.0%; Score 47; DB 2; Length 482;
Best Local Similarity 66.7%; Pred. No. 5.2;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 DFTNHLEDTRRNINK 19
Db 347 DFTLHLEDGSYNIRK 361

RESULT 12
US-08-135-511-29
Sequence 29, Application US/08135511
Patent No. 5558999
GENERAL INFORMATION:
APPLICANT: Chiang, John
TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene
TITLE OF INVENTION: Regulatory Elements and Methods for Using Them
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/135,511
FILING DATE: 13-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SANDERCOCK, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 18748/175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-135-511-29

Query Match 47.0%; Score 47; DB 1; Length 503;
Best Local Similarity 66.7%; Pred. No. 5.5;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 DFTNHLEDTRRNINK 19
Db 369 DFTLHLEDGSYNIRK 383

RESULT 13
US-08-187-453-29
Sequence 29, Application US/08187453
Patent No. 5753431
GENERAL INFORMATION:
APPLICANT: Chiang, John
TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene
TITLE OF INVENTION: Regulatory Elements and Transcription Factors
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/187,453
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/135,488
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/135,511
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/135,510
FILING DATE: 13-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: SANDERCOCK, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 18748/188
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-187-453-29

Query Match 47.0%; Score 47; DB 1; Length 503;
Best Local Similarity 66.7%; Pred. No. 5.5;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 DFTNHLEDTRRNINK 19

DB 369 DFTLHLEDGSYNIRK 383

RESULT 14

US-08-135-511-28

Sequence 28, Application US/08135511

Patent No. 5558999

GENERAL INFORMATION:

APPLICANT: Chiang, John

TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene

TITLE OF INVENTION: Regulatory Elements and Methods for Using Them

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08-135,511

FILING DATE: 13-OCT-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SANDERCOCK, Colin G.

REGISTRATION NUMBER: 31,298

REFERENCE/DOCKET NUMBER: 18748/175

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 504 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-135-511-28

Query Match 47.0%; Score 47; DB 1; Length 504;
Best Local Similarity 66.7%; Pred. No. 5.5;

Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 DFTNHLLEDTRRNIRK 19
DB 369 DFTLHLEDGSYNIRK 383

RESULT 15

US-08-135-511-30

Sequence 30, Application US/08135511

Patent No. 5558999

GENERAL INFORMATION:

APPLICANT: Chiang, John

TITLE OF INVENTION: Cholesterol 7a-Hydroxylase Gene

TITLE OF INVENTION: Regulatory Elements and Methods for Using Them

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/135,511

FILING DATE: 13-OCT-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SANDERCOCK, Colin G.

REGISTRATION NUMBER: 31,298

REFERENCE/DOCKET NUMBER: 18748/175

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 504 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-135-511-30

Query Match 47.0%; Score 47; DB 1; Length 504;
Best Local Similarity 66.7%; Pred. No. 5.5;

Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 5 DFTNHLLEDTRRNIRK 19
DB 369 DFTLHLEDGSYNIRK 383

Search completed: July 11, 2003, 12:02:06
Job time : 4.39709 secs

THIS PAGE BLANK (USP10)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:59:17 ; Search time 5.56965 Seconds
(Without alignments)
397.148 Million cell updates/sec

Title: US-10-091-442-3
Perfect score: 100
Sequence: 1 VERVDFTHLEDTRRNINK 19

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Published_Applications_AA:*
2: /cgn2_6/ptodata/1/pubppaa/US07_NEM_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/PCF_NEM_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_NEM_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/PCFUS_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEM_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	19	9 US-10-091-442-3	Sequence 3, Appl1
2	100	100.0	19	9 US-09-140-719-3	Sequence 34, Appl1
3	100	100.0	380	9 US-10-091-442-34	Sequence 34, Appl1
4	100	100.0	380	10 US-09-140-719-34	Sequence 34, Appl1
5	50	50.0	390	7 US-08-731-566-2	Sequence 2, Appl1
6	50	50.0	390	9 US-09-823-187-29	Sequence 29, Appl1
7	50	50.0	390	9 US-09-823-187-30	Sequence 29, Appl1
8	46	46.0	390	9 US-09-823-187-27	Sequence 27, Appl1
9	46	46.0	390	9 US-09-823-187-28	Sequence 28, Appl1
10	46	46.0	410	9 US-09-910-186A-32	Sequence 32, Appl1
11	45	45.0	45	10 US-09-864-761-41911	Sequence 41911, A
12	45	45.0	123	9 US-10-073-979-21	Sequence 21, Appl1
13	45	45.0	123	9 US-09-764-891-4883	Sequence 4883, Ap
14	45	45.0	123	10 US-09-764-898-242	Sequence 242, App
15	45	45.0	391	9 US-10-157-031-355	Sequence 355, App
16	45	45.0	391	9 US-09-466-386A-110	Sequence 110, App
17	45	45.0	391	10 US-09-735-705-110	Sequence 110, App
18	45	45.0	391	10 US-09-850-716A-110	Sequence 110, App
19	45	45.0	391	10 US-09-897-778-110	Sequence 110, App

20	45	45.0	400	9 US-09-466-396A-112	Sequence 112, App
21	45	45.0	400	10 US-09-735-705-112	Sequence 112, App
22	45	45.0	400	10 US-09-850-716A-112	Sequence 112, App
23	45	45.0	400	10 US-09-897-778-112	Sequence 112, App
24	45	45.0	612	10 US-09-925-301-1358	Sequence 1358, Ap
25	44	44.0	340	10 US-09-764-898-169	Sequence 169, App
26	44	44.0	392	9 US-10-198-070-62	Sequence 62, Appl
27	44	44.0	617	9 US-10-012-896-947	Sequence 947, App
28	44	44.0	617	9 US-09-895-814-947	Sequence 947, App
29	42	42.0	379	9 US-10-012-896-946	Sequence 946, App
30	42	42.0	379	9 US-09-895-814-946	Sequence 946, App
31	42	42.0	392	9 US-10-198-070-67	Sequence 67, Appl
32	42	42.0	392	9 US-10-198-070-70	Sequence 70, Appl
33	41	41.0	395	9 US-09-823-187-2	Sequence 2, Appl1
34	41	41.0	437	10 US-09-925-301-1200	Sequence 1200, Ap
35	41	41.0	459	10 US-09-925-300-1440	Sequence 1440, Ap
36	41	41.0	531	9 US-10-156-761-14639	Sequence 14639, A
37	41	41.0	558	9 US-09-941-831-28	Sequence 28, Appl
38	41	41.0	705	9 US-10-128-714-3201	Sequence 3201, Ap
39	41	41.0	748	9 US-10-128-714-8201	Sequence 8201, Ap
40	41	41.0	818	9 US-09-738-626-6491	Sequence 6491, Ap
41	40.5	40.5	1019	10 US-09-801-574-76	Sequence 76, Appl
42	40	40.0	150	9 US-10-001-857-203	Sequence 203, Appl
43	40	40.0	433	9 US-09-738-626-6592	Sequence 6592, Ap
44	39.5	39.5	745	9 US-10-000-823-4	Sequence 4, Appl1
45	39.5	39.5	761	9 US-10-011-582-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-091-442-3
Sequence 3, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
IWASA, Fuyuki
TSURUOKA, No. US20020164711A1uo
NAKAZATO, Hiroshi
MIDURA, Kenju
ISHIDA, No. US20020164711A1uhhro
KURIHARA, Tatsuya
YAMACHI, Kozo
YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091.442
FILING DATE: 07-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140.719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474.661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091.028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339

FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-091-442-3

Query Match 100.0%; Score 100; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERDYFTNHLIEDTRRNINK 19
DB 1 VERDYFTNHLIEDTRRNINK 19

RESULT 2
US-09-140-719-3
Sequence 3, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURODOKA, No. US20010026931A1uo
APPLICANT: MAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20010026931A1uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20010026931A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
State: Virginia
Country: United States
Zip: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.

REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-140-719-3

Query Match 100.0%; Score 100; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERDYFTNHLIEDTRRNINK 19
DB 1 VERDYFTNHLIEDTRRNINK 19

RESULT 3
US-10-091-442-34
Sequence 34, Application US/10091442
Patent No. US20020164711A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masafumi
APPLICANT: IWASA, Fuyuki
APPLICANT: TSURODOKA, No. US20020164711A1uo
APPLICANT: MAKAZATO, Hiroshi
APPLICANT: MIURA, Kenju
APPLICANT: ISHIDA, No. US20020164711A1uhiro
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAICHI, Kozo
APPLICANT: YAMAGUCHI, No. US20020164711A1omi
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
City: Alexandria
State: Virginia
Country: United States
Zip: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/091,442
FILING DATE: 07-Mar-2002.
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-10-091-442-34

Query Match 100.0%; Score 100; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 2,1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTHLEDTRNIN 19
DB 126 VERVDFTHLEDTRNIN 144

RESULT 4
US-09-140-719-34
Sequence 34, Application US/09140719
Patent No. US20010026931A1
GENERAL INFORMATION:
APPLICANT: TSUJIMOTO, Masaaki
APPLICANT: IWASHI, Fuyuki
APPLICANT: TSURUOKA, No. US20010026931A1
APPLICANT: NAKAZATO, Hiroshi
APPLICANT: MIURA, Kenji
APPLICANT: ISHIDA, No. US20010026931A1
APPLICANT: KURIHARA, Tatsuya
APPLICANT: YAMAGUCHI, Kozo
TITLE OF INVENTION: MEGAKARYOCYTE DIFFERENTIATION FACTOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,719
FILING DATE: 08-AUG-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,661
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/091,028
FILING DATE: 14-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-212305
FILING DATE: 17-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-067339
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 001560-247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-140-719-34

Query Match 100.0%; Score 100; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 2,1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VERVDFTHLEDTRNIN 19
DB 126 VERVDFTHLEDTRNIN 144

RESULT 5
US-08-731-566-2
Sequence 2, Application US/08731566
Publication No. US20030072752A1
GENERAL INFORMATION:
APPLICANT: Toshihiko Hibino et al.
TITLE OF INVENTION: TWO NOVEL GENES FROM PSORIATIC EPIDERMIS:
PSORIASIS TYPE I AND PSORIASIS TYPE II
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1775
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,566
FILING DATE: 16-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,755
FILING DATE: 20-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,679
FILING DATE: 17-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-040CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 390 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-731-566-2

Query Match 50.0%; Score 50; DB 7; Length 390;
Best Local Similarity 55.6%; Pred. No. 3;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 VERVDFTHLEDTRNIN 18
DB 131 VESVDFANAPESRRKIN 148

RESULT 6
US-09-823-187-29
Sequence 29, Application US/09823187
Publication No. US20030096952A1
GENERAL INFORMATION:
APPLICANT: Burgess, Catherine

```

; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigar, Muralidhar
; APPLICANT: Patnrajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-29

Query Match      50.0%; Score 50; DB 9; Length 390;
Best Local Similarity 55.6%; Pred. No. 3;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```

QY      1 VERVDFTNHLDETRRNIN 18
DB      131 VESVDPRNAPESRKRIN 148

RESULT 7
US-09-823-187-30
; Sequence 30, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigar, Muralidhar
; APPLICANT: Patnrajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
```

```

; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-30

Query Match      50.0%; Score 50; DB 9; Length 390;
Best Local Similarity 55.6%; Pred. No. 3;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```

QY      1 VERVDFTNHLDETRRNIN 18
DB      131 VESVDPRNAPESRKRIN 148

RESULT 8
US-09-823-187-27
; Sequence 27, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigar, Muralidhar
; APPLICANT: Patnrajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
```



```
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41911
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023009.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: SWISSPROT HIT: P00488, EVALUATION: 8.00e-21
; OTHER INFORMATION: EST_HUMAN HIT: AV733491.1, EVALUATION: 3.00e-15
; US-09-864-761-41911

Query Match      45.0%; Score 45; DB 10; Length 45;
Best Local Similarity 57.1%; Pred. No. 1.7;
Matches      8; Conservative      4; Mismatches      2; Indels      0; Gaps      0;

QY      4 VERDFTNHLDETRRNI 17
DB      15 VERDFTNHLDETRRNI 28

RESULT 12
US-10-073-979-21
; Sequence 21, Application US/10073979
; Publication No. US20030044905A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT231C1
; CURRENT APPLICATION NUMBER: US/10/073,979
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (88)
; OTHER INFORMATION: (88)
```

```
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-073-979-21

Query Match      45.0%; Score 45; DB 9; Length 123;
Best Local Similarity 44.4%; Pred. No. 5.3;
Matches      8; Conservative      4; Mismatches      6; Indels      0; Gaps      0;

QY      1 VERDFTNHLDETRRNI 18
DB      82 LQTVDFXQSTEEFTRKTI 99

RESULT 13
US-09-764-891-4883
; Sequence 4883, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4883
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-891-4883

Query Match      45.0%; Score 45; DB 9; Length 123;
Best Local Similarity 44.4%; Pred. No. 5.3;
Matches      8; Conservative      4; Mismatches      6; Indels      0; Gaps      0;

QY      1 VERDFTNHLDETRRNI 18
DB      82 LQTVDFXQSTEEFTRKTI 99

RESULT 14
US-09-764-898-242
; Sequence 242, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 242
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-898-242

Query Match      45.0%; Score 45; DB 10; Length 123;
Best Local Similarity 44.4%; Pred. No. 5.3;
Matches      8; Conservative      4; Mismatches      6; Indels      0; Gaps      0;

QY      1 VERDFTNHLDETRRNI 18
DB      82 LQTVDFXQSTEEFTRKTI 99
```


Db 82 LQVDFXQSTEEETKRTIN 99

RESULT 15

US-10-157-031-355

; Sequence 355, Application US/10157031
; Publication No. US20030108890A1

GENERAL INFORMATION:

; APPLICANT: Baranova, A. V.

; APPLICANT: Yankovsky, N. K.

; APPLICANT: Kozlov, A. P.

; APPLICANT: Lodashev, A. V.

; APPLICANT: Krutovskaya, L. L.

; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences

; FILE REFERENCE: 2760-103

; CURRENT APPLICATION NUMBER: US/10/157,031

; CURRENT FILING DATE: 2002-05-30

; NUMBER OF SEQ ID NOS: 415

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 355

; LENGTH: 391

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-157-031-355

Query Match 45.0%; Score 45; DB 9; Length 391;

Best Local Similarity 44.4%; Pred. No. 20;

Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 VERVDFTHLEDTRRIN 18

Db 132 LQVDFVNADESRKKIN 149

Search completed: July 11, 2003, 12:37:32
Job time : 5.56965 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 11, 2003, 11:45:06 ; Search time 3.8316 Seconds
(without alignments)
476.708 Million cell updates/sec

Title: US-10-091-442-3

Perfect score: 100

Sequence: 1 VERVDFTNHLEDTFRNINK.19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	50.0	390	2	I38201
2	49	49.0	1169	2	H70178
3	48	48.0	472	2	T27755
4	47	47.0	501	2	I46701
5	47	47.0	503	1	A35376
6	47	47.0	503	2	A54779
7	47	47.0	504	1	JH0659
8	47	47.0	504	2	S39399
9	47	47.0	1046	2	T42720
10	47	47.0	1046	2	T42734
11	46	46.0	375	2	A36898
12	46	46.0	390	2	I38202
13	46	46.0	1287	2	S39791
14	46	46.0	2022	2	T43214
15	45.5	45.5	163	2	D64427
16	45	45.0	391	2	JC7118
17	45	45.0	673	1	B64530
18	45	45.0	673	2	A71978
19	44.5	44.5	330	2	D97235
20	44.5	44.5	333	2	S57050
21	44	44.0	348	2	E82933
22	44	44.0	369	2	T48612
23	44	44.0	483	2	JC5743
24	44	44.0	555	1	SC5HMA
25	44	44.0	1147	2	S64830
26	43.5	43.5	119	2	S76276
27	43	43.0	58	2	G95132
28	43	43.0	259	2	G81833
29	43	43.0	301	2	C64633

30	43	43.0	363	2	F71881	probable flagellar
31	43	43.0	468	2	F84686	hypothetical prote
32	43	43.0	732	1	EKHUX	protein-glutamine
33	43	43.0	1111	2	A86922	probable arabinosy
34	43	43.0	1461	2	T26327	hypothetical prote
35	43	43.0	1619	2	T18499	hypothetical prote
36	42.5	42.5	409	2	A87495	hypothetical prote
37	42	42.0	399	2	E71188	hypothetical prote
38	42	42.0	589	2	S71954	RNA/ssDNA-binding
39	42	42.0	589	2	S23991	conjugative transf
40	42	42.0	957	2	H69141	hypothetical prote
41	41.5	41.5	349	2	AH1764	B. subtilis probab
42	41	41.0	288	2	AD0884	probable transfer
43	41	41.0	317	2	T27330	hypothetical prote
44	41	41.0	319	2	F96930	molybdate-binding
45	41	41.0	347	2	S44846	K06H7.6 protein -

ALIGNMENTS

RESULT 1
I38201
Squamous cell carcinoma antigen.1 - human
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence: revision 23-Feb-1996 #text_change 26-May-2000
C:Accession: I38201: I38200: G01631: J70966: J70967
R:Schneider, S.S.; Schick, C.; Fish, K.E.; Miller, E.; Pena, J.C.; Treter, S.D.; Hui,
Proc. Natl. Acad. Sci. U.S.A. 92, 3147-3151, 1995
A:Title: A serine proteinase inhibitor locus at 16q21.3 contains a tandem duplication
A:Reference number: I38200; M01D:95241462; PMID:7724531
A:Accession: I38201
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-390 <SCH1>
A:Cross-references: EMBL:U19568; GB:U19558; NID:q1172085
A:Accession: I38200
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-117 <SCH2>
A:Cross-references: EMBL:U19562; NID:g897835; PIDN:AAA86316.1; PID:g897844
R:Silverman, G.A.
Submitted to the EMBL Data Library, January 1995
A:Reference number: G07968
A:Accession: G01631
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-350, 'G', 352-390 <SIL>
A:Cross-references: EMBL:U19556; NID:q1276435; PID:q1052869
R:Sumitani, Y.; Kishi, F.; Sekiguchi, K.; Kato, H.
Biochem. Biophys. Res. Commun. 181, 51-58, 1991
A:Title: Squamous cell carcinoma antigen is a new member of the serine protease inhib
A:Reference number: J70966; M01D:92068241; PMID:1958219
A:Accession: J70966
A:Molecule type: mRNA
A:Residues: 1-350, 'G', 352-356, 'A', 358-390 <SUM1>
A:Cross-references: GB:S66896; NID:q239551; PIDN:AA820405.1; PID:q239552
A:Accession: J70967
A:Molecule type: protein
A:Residues: 11-21; 231-237; 240-256; 303-325 <SUM2>
C:Comment: This antigen probably acts as a proteinase inhibitor to modulate the host
C:Genetics:
A:Gene: GDB:SCCA1: SCC
A:Cross-references: GDB:625364; OMIM:600517
A:Map position: 18q21.3-18q21.3
A:Introns: 55/3; 74/3; 117/3; 157/1; 204/3; 256/3
C:Superfamily: antithrombin III
C:Keywords: cysteine proteinase inhibitor; glycoprotein
F:65,93,171,376/binding site: carboxylate (asn) (covalent) #status predicted
F:334/inhibitory site: Ser (cathepsin L) #status predicted

Query Match 50.0%; Score 50; DB 2; Length 390;
Best Local Similarity 55.6%; Pred. No. 3.3;

Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VERVDFTNHLEDFRRNIN 18
 |||||
 Db 131 VESVDFNAPEESRRKIN 148

RESULT 2

H70178
 exodeoxyribonuclease V, beta chain (recB) homolog - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C:Accession: H70178
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Iatnigra, R.; White
 son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98055943; PMID:9403685
 A:Accession: H70178
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1169 <KLE>
 A:Cross-references: GB:AE001164; GB:AE000783; NID:92688550; PIDN:AAC66981.1; PID:9268855
 A:Experimental source: strain B31

Query Match 49.0%; Score 49; DB 2; Length 1169;
 Best Local Similarity 42.1%; Pred. No. 16;
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 VERVDFTNHLEDFRRNIN 19
 |||||
 Db 459 IEKIDFTNSLPNOKNDNNK 477

RESULT 3

T27755
 hypothetical protein ZK1320.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27755
 R:Berts, M.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: Z20414
 A:Accession: T27755
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-472 <WIL>
 A:Cross-references: EMBL:Z46934; PIDN:CAAB7047.1; GSPDB:GN00020; CESP:ZK1320.9
 A:Experimental source: clone ZK1320
 C:Genetics:
 A:Gene: CESP:ZK1320.9
 A:Map position: 2
 A:Introns: 19/2; 55/1; 106/1; 160/1; 186/1; 323/1; 411/3

Query Match 48.0%; Score 48; DB 2; Length 472;
 Best Local Similarity 66.7%; Pred. No. 8.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 DFTNHLEDFRRN 16
 |||||
 Db 328 DFTNHIDIVRRN 339

RESULT 4

I46701
 cholesterol 7alpha-monooxygenase (EC 1.14.13.17) - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 04-Mar-2000
 C:Accession: I46701
 R:Kai, M.; Eto, T.; Kondo, K.; Setoguchi, Y.; Higashi, S.; Maeda, Y.; Setoguchi, T.
 J. Lipid Res. 36, 367-374, 1995

A:Title: Synchronous circadian rhythms of mRNA levels and activities of cholesterol 7
 A:Reference number: I46701; MUID:95271192; PMID:7751825
 A:Accession: I46701

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-501 <KAI>

A:Cross-references: GB:U10754; NID:9437659; PIDN:AAV4382.1; PID:9437660
 C:Genetics:
 A:Gene: CYP7
 C:Superfamily: human cytochrome P450 CYP7A1; cytochrome P450 homology
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase; transmembrane
 F:279-463/Domain: cytochrome P450 homology <P45>
 F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.0%; Score 47; DB 2; Length 501;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 DFTNHLEDFRRNIN 19
 |||||
 Db 366 DFTNHLEDFGYNIRK 380

RESULT 5

A33376
 cholesterol 7alpha-monooxygenase (EC 1.14.13.17) - rat
 N:Alternate names: cholesterol 7alpha-hydroxylase
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 21-Jul-2000
 C:Accession: A33376; A35609; A36450; S06632; A37011; A38736; S27206
 R:Jelinek, D.F.; Andersson, S.; Staughter, C.A.; Russell, D.W.
 J. Biol. Chem. 265, 8190-8197, 1990
 A:Title: Cloning and regulation of cholesterol 7alpha-hydroxylase, the rate-limiting
 A:Reference number: A33376; MUID:90243699; PMID:2335522
 A:Accession: A33376

A:Molecule type: mRNA
 A:Residues: 1-503 <JEL>

A:Cross-references: GB:J05430; NID:9203792; PIDN:AAA41041.1; PID:9203793
 R:Jelinek, D.F.; Russell, D.W.
 Biochemistry 29, 7781-7785, 1990

A:Title: Structure of the rat gene encoding cholesterol 7alpha-hydroxylase.
 A:Reference number: A35609; MUID:91084435; PMID:2261433

A:Accession: A35609
 A:Molecule type: DNA

A:Residues: 1-26 <JEL2>
 A:Cross-references: GB:J02926
 R:Noshiro, M.; Nishimoto, M.; Okuda, K.

J. Biol. Chem. 265, 10036-10041, 1990
 A:Title: Rat liver cholesterol 7alpha-hydroxylase. Pretranslational regulation for cl
 A:Reference number: A36450; MUID:90277612; PMID:1693613

A:Accession: A36450
 A:Molecule type: mRNA

A:Residues: 1-503 <NOS1>
 A:Cross-references: GB:J05460; NID:9203455; PIDN:AAA03649.1; PID:9203456
 R:Noshiro, M.; Nishimoto, M.; Morohashi, K.I.; Okuda, K.

FEBS Lett. 257, 97-100, 1989
 A:Title: Molecular cloning of cDNA for cholesterol 7alpha-hydroxylase from rat liver
 A:Reference number: S06632; MUID:90033362; PMID:2806567

A:Accession: S06632
 A:Molecule type: mRNA

A:Residues: 1-503 <NOS2>
 A:Cross-references: GB:X17595; NID:957535; PIDN:CAB57878.1; PID:96018697
 R:Li, Y.C.; Wang, D.P.; Chiang, J.Y.L.

J. Biol. Chem. 265, 12012-12019, 1990
 A:Title: Regulation of cholesterol 7alpha-hydroxylase in the liver. Cloning, sequenci
 A:Reference number: A37071; MUID:90307735; PMID:1694652

A:Accession: A37071
 A:Molecule type: mRNA

A:Residues: 1-503 <LIA>
 A:Cross-references: GB:J05509; NID:9203204; PIDN:AAA0899.1; PID:9203205
 R:Nishimoto, M.; Gotch, O.; Okuda, K.; Noshiro, M.

J. Biol. Chem. 266, 6467-6471, 1991
 A:Title: Structural analysis of the gene encoding rat cholesterol alpha-hydroxylase.

A:Reference number: A37071; MUID:90307735; PMID:1694652

A:Reference number: A38736; MUID:91177904; PMID:2007596
A:Accession: A38736
A:Molecule type: DNA
A:Residues: 1-370, 'S', 372-503 <NIS>
A:CROSS-references: GB:M59184
R:Chiang, J.Y.L.; Yang, T.P.; Wang, D.P.
Biochem. Biophys. Acta 1132, 337-339, 1992
A:Title: Cloning and 5'-flanking sequence of a rat cholesterol 7alpha-hydroxylase gene.
A:Reference number: S27206; MUID:93041942; PMID:1420318
A:Accession: S27206
A:Molecule type: DNA
A:Residues: 1-26 <CHT>
A:CROSS-references: GB:214108; GB:S48135; NID:955835; PIDN:CAA78481.1; PID:955836
C:Comment: This liver microsomal enzyme catalyzes the conversion of cholesterol to bile
C:Genetics:
A:Introns: 27/2; 107/3; 303/2; 347/1; 405/3
C:Superfamily: human cytochrome P450 CYP7A1; cytochrome P450 homology
C:Keywords: cholesterol metabolism; chromoprotein; electron transfer; heme; iron; metal
F:282-466/Domain: cytochrome P450 homology <P45>
F:444/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.0%; Score 47; DB 1; Length 503;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 5 DFTNHLDETRRNK 19
||| |||| |||
Db 369 DFTLHLEDGSYNIRK 383

RESULT 6
A54779
cholesterol 7alpha-monooxygenase (EC 1.14.13.17) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 04-Mar-2000
C:Accession: A54779
R:Tung, K.W.; Ishimura-Oka, K.; Kihara, S.; Oka, K.; Chan, L.
Genomics 21, 244-247, 1994
A:Title: Structure of the mouse cholesterol 7alpha-hydroxylase gene.
A:Reference number: A54779; MUID:94375022; PMID:8088795
A:Accession: A54779
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <TRU>
A:CROSS-references: GB:L23754
C:Genetics:
A:Gene: CYP7
C:Superfamily: human cytochrome P450 CYP7A1; cytochrome P450 homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase; transmembrane pro
F:282-466/Domain: cytochrome P450 homology <P45>
F:444/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.0%; Score 47; DB 2; Length 503;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 5 DFTNHLDETRRNK 19
||| |||| |||
Db 369 DFTLHLEDGSYNIRK 383

RESULT 7
JH0659
cholesterol 7alpha-monooxygenase (EC 1.14.13.17) chain 1 - human
N:Alternate names: cholesterol 7alpha-hydroxylase; cytochrome P450, subfamily VIIA
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 03-Mar-2000
C:Accession: S29818; JH0659; S11051; A42201; A54310
R:Nishimoto, M.; Noshiro, M.; Okuda, K.
Biochim. Biophys. Acta 1172, 147-150, 1993
A:Title: Structure of the gene encoding human liver cholesterol 7alpha-hydroxylase.
A:Reference number: S29818; MUID:93176797; PMID:8439551
A:Accession: S29818

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-504 <NIS>
A:CROSS-references: EMBL:L04633
R:Karam, W.G.; Chiang, J.Y.L.
Biochem. Biophys. Res. Commun. 185, 588-595, 1992
A:Title: Polymorphisms of human cholesterol 7 alpha-hydroxylase.
A:Reference number: JH0659; MUID:92304280; PMID:1610352
A:Accession: JH0659

A:Molecule type: mRNA
A:Residues: 1-504 <KAR>
A:CROSS-references: GB:M93133; NID:9181318; PIDN:AAA58435.1; PID:9181319
A:Experimental source: liver
A:Note: 100-ser was also found
R:Noshiro, M.; Okuda, K.
FEBS Lett. 268, 137-140, 1990
A:Title: Molecular cloning and sequence analysis of cDNA encoding human cholesterol 7
A:Reference number: S11051; MUID:90346120; PMID:2384150
A:Accession: S11051
A:Molecule type: mRNA
A:Residues: 1-346, 'N', 348-384, 'S', 386-504 <NOS>
A:CROSS-references: GB:M93133
R:Molova, D.T.; Chen, W.S.; Cimals, G.M.; Tan, C.P.
Biochemistry 31, 2539-2544, 1992
A:Title: Transcriptional regulation of the human cholesterol 7 alpha-hydroxylase gene
A:Reference number: A42201; MUID:92190183; PMID:1312351
A:Accession: A42201

A:Molecule type: DNA
A:Residues: 1-25 <MOU>
A:CROSS-references: GB:M09647; GB:J05363; NID:9180469; PIDN:AAA58423.1; PID:9553228
A:Note: Sequence extracted from NCBI backbone (NCBIN:89078, NCBI:89079)
R:Wang, D.P.; Chiang, J.Y.
Genomics 20, 320-323, 1994
A:Title: Structure and nucleotide sequences of the human cholesterol 7 alpha-hydroxyl
A:Reference number: A54310; MUID:94292222; PMID:8020987
A:Accession: A54310
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-140 <RES>
A:CROSS-references: GB:L13460; NID:9489855; PIDN:AAA61350.1; PID:9624966
C:Comment: This liver microsomal enzyme catalyzes the conversion of cholesterol to bl
C:Genetics:
A:Gene: GDB:CYP7A1; CYP7
A:CROSS-references: GDB:132221; OMIM:118455
A:Map position: 8q11-8q12
A:Introns: 27/2; 107/3; 303/2; 347/1; 405/3
C:Superfamily: human cytochrome P450 CYP7A1; cytochrome P450 homology
C:Keywords: cholesterol metabolism; chromoprotein; electron transfer; heme; iron; met
F:282-466/Domain: cytochrome P450 homology <CYP>
F:444/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.0%; Score 47; DB 1; Length 504;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 5 DFTNHLDETRRNK 19
||| |||| |||
Db 369 DFTLHLEDGSYNIRK 383

RESULT 8
S39399
cholesterol 7alpha-monooxygenase (EC 1.14.13.17) - Chinese hamster
N:Alternate names: cholesterol 7alpha-hydroxylase
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 04-Mar-2000
C:Accession: S39399
R:Crestani, M.; Galli, G.; Chiang, J.Y.L.
Arch. Biochem. Biophys. 306, 451-460, 1993
A:Title: Genomic cloning, sequencing, and analysis of the hamster cholesterol 7-alpha
A:Reference number: S39399; MUID:94029018; PMID:8105753
A:Accession: S39399
A:Status: preliminary

C;Keywords: glycoprotein; serine proteinase inhibitor
F;1-22/Domain: signal sequence #status predicted <SIG>

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2003, 11:43:41 ; Search time 1.81705 Seconds
(without alignments)
433.698 Million cell updates/sec

Title: US-10-091-442-3

Perfect score: 100

Sequence: 1 VERVDFTHLEDTRRNINK 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	380	SPB7_HUMAN	O75635 homo sapien
2	50	50.0	1	SCC1_HUMAN	P29508 homo sapien
3	49	49.0	1169	EX5B_BORBU	O51578 borrelia bu
4	48.5	48.5	210	NOC4_HUMAN	O43402 homo sapien
5	47.5	47.5	207	NOC4_MOUSE	O70378 mus musculi
6	47	47.0	501	CP7A_MOUSE	P51542 oryctolagus
7	47	47.0	503	CP7A_RABIT	O64505 mus musculi
8	47	47.0	503	CP7A_RAT	P18125 rattus norv
9	47	47.0	504	CP7A_CRIGR	P46634 cricetulus
10	47	47.0	504	CP7A_HUMAN	P22680 homo sapien
11	46	46.0	375	MASP_HUMAN	P36952 homo sapien
12	46	46.0	375	MASP_MOUSE	P70124 mus musculi
13	46	46.0	375	MASP_RAT	P70564 rattus norv
14	46	46.0	390	SCC2_HUMAN	O60393 clostridium
15	46	46.0	1296	BXG_CLOBO	O60393 clostridium
16	46	46.0	2022	ANT1_ONCYO	O58427 methanococ
17	45.5	45.5	163	YA21_METJA	O94108 homo sapien
18	45	45.0	391	SB13_HUMAN	O94108 homo sapien
19	44.5	44.5	483	CP7A_YEAST	P47103 saccharomyc
20	44	44.0	501	MM20_PIG	P79877 sus scrofa
21	44	44.0	501	CP7A_PIG	P21360 pichia angu
22	44	44.0	555	MASY_PICAN	P13386 saccharomyc
23	44	44.0	1147	KIN2_YEAST	O94108 homo sapien
24	43	43.0	505	CSC1_DROME	P00488 homo sapien
25	43	43.0	731	F13A_HUMAN	O96615 mycobacteri
26	43	43.0	1111	EMBA_MYCLE	O96615 mycobacteri
27	42	42.0	392	SB11_HUMAN	O96615 mycobacteri
28	42	42.0	592	R856_HUMAN	O96615 mycobacteri
29	42	42.0	602	TRAN_ECOLI	P24082 escherichia
30	41	41.0	203	U172_DROME	O94108 homo sapien
31	41	41.0	215	HM00_CORDI	P71191 dirosophila
32	41	41.0	285	STAR_BRARE	O94910 brachydanio
33	41	41.0	376	PT16_HUMAN	P35237 homo sapien

34	41	41.0	378	1	PT16_BOVIN	O02739 bos taurus
35	41	41.0	409	1	ILEU_HUMAN	P30740 homo sapien
36	41	41.0	475	1	SB12_HUMAN	O96615 mycobacteri
37	41	41.0	728	1	YMK6_CAEEL	P34514 caenorhabdi
38	41	41.0	1161	1	DP3A_EUCAL	P57332 buchnera ap
39	41	41.0	1995	1	YCX7_CHLRE	P36495 chlamydomon
40	41	41.0	2364	1	SPCO_HUMAN	O01082 homo sapien
41	40.5	40.5	578	1	YDEM_CAEEL	O19124 caenorhabdi
42	40.5	40.5	842	1	SYA_CAMEL	O9p105 campylobact
43	40.5	40.5	1019	1	STJ1_HUMAN	O9p105 campylobact
44	40	40.0	171	1	VPE2_BPP2	P22502 bacteriopia
45	40	40.0	289	1	STX3_HUMAN	O13277 homo sapien

ALIGNMENTS

RESULT 1
ID SPB7_HUMAN STANDARD: PRT; 380 AA.
AC O75635:
DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Megsin (TP55) (Serpin B7).
GN SERPINB7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97326116; PubMed-9182567;
TSUJIMOTO M., TSURUOKA N., ISHIDA N., KURIHARA T., IWASA F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Hanyama M., Miura K.,
RA Nakajishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity".
RL J. Biol. Chem. 272:15373-15380(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-98376492; PubMed-9710452;
TSUJIMOTO M., TSURUOKA N., ISHIDA N., KURIHARA T., IWASA F.,
RA Yamashiro K., Rogi T., Kodama S., Katsuragi N., Adachi M.,
RA Katayama T., Nakao M., Yamachi K., Hashino J., Hanyama M., Miura K.,
RA Nakajishi T., Nakazato H., Teramura M., Mizoguchi H., Yamaguchi N.,
RT "Purification, cDNA cloning, and characterization of a new serpin with
RT megakaryocyte maturation activity".
RL J. Biol. Chem. 272:15373-15380(1997).
CC -!- FUNCTION: Might function as an inhibitor of lys-specific
CC proteases. Might influence the maturation of megakaryocytes via
CC its action as a serpin.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Predominantly expressed in mesangial cells.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D88575; BAA1232.1; -.
CC EMBL: AF027866; AAC64506.1; -.
CC HSSP: P05619; 1HLE.
CC Gene: HGNC:13902; SERPINB7.
CC MIM: 603357; -.
CC InterPro: IPR000215; Serpin.
CC Pfam: PF00079; serpin; 1.
CC SMART: SM00093; SERPIN; 1.
CC PROSITE: PS00284; SERPIN; 1.

KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 347 348 REACTIVE BOND (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 42904 MW; 9A2CDB6C63CFF605 CRC64;
 Query Match 100.0%; Score 100; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 1.2e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VERVDFNHLDEFRNRINK 19
 |||
 Db 126 VERVDFNHLDEFRNRINK 144

RESULT 2
 SCCL_HUMAN
 ID SCCL_HUMAN STANDARD; PRT; 390 AA.
 AC P29508: 096J21;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Squamous cell carcinoma antigen 1 (SCCA-1) (Protein T4-A).
 GN SERPINB3 OR SCCA1 OR SCCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=92068241; PubMed=1958219;
 RA Sumnaml Y., Kishi F., Sekiguchi K., Kato H.;
 RT "Squamous cell carcinoma antigen is a new member of the serine
 protease inhibitors";
 RL Biochem. Biophys. Res. Commun. 181:51-58(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95241462; PubMed=7724531;
 RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,
 RA Treter S.D., Hul S.M., Silverman G.A.;
 RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem
 duplication of the human squamous cell carcinoma antigen gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
 IMMUNE RESPONSE AGAINST TUMOR CELLS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.
 CC -1- DEVELOPMENTAL STAGE: ITS EXPRESSION IS CLOSELY RELATED TO CELLULAR
 DIFFERENTIATION IN BOTH NORMAL AND MALIGNANT SQUAMOUS CELLS.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; S66896; AAB20405.1; -;
 DR EMBL; U19556; AAA97552.1; -;
 DR EMBL; U19568; AAA86317.1; -;
 DR EMBL; U19559; AAA86317.1; JOINED.
 DR EMBL; U19560; AAA86317.1; JOINED.
 DR EMBL; U19562; AAA86317.1; JOINED.
 DR EMBL; U19565; AAA86317.1; JOINED.
 DR EMBL; U19567; AAA86317.1; JOINED.
 DR EMBL; U19562; AAA86316.1; -;
 DR EMBL; U19559; AAA86316.1; JOINED.
 DR EMBL; U19560; AAA86316.1; JOINED.
 DR EMBL; U19560; AAA86316.1; JOINED.

DR EMBL; BC005224; AAH05224.1; -;
 DR PIR; J10966; J10966.
 DR HSSP; P01008; IATH.
 DR Genew; HGNC:10569; SERPINB3.
 DR MIM; 600517; -;
 DR InterPro; IPR00215; Serpin.
 DR Pfam; PF00079; serpin.1.
 DR SMART; SM00093; SERPIN.1.
 DR PROSITE; PS00284; SERPIN.1.
 KW Serpin; Serine protease inhibitor.
 FT ACT_SITE 354 355 REACTIVE BOND.
 FT CONFLICT 357 357 T -> A (IN REF. 1).
 SQ SEQUENCE 390 AA; 44564 MW; E5F27F986C752CFA CRC64;
 Query Match 50.0%; Score 50; DB 1; Length 390;
 Best Local Similarity 55.6%; Pred. No. 1.2;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VERVDFNHLDEFRNRIN 18
 |||
 Db 131 VERVDFNHLDEFRNRIN 148

RESULT 3
 EX5B_BORBU
 ID EX5B_BORBU STANDARD; PRT; 1169 AA.
 AC 051578;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exodeoxyribonuclease V beta chain (EC 3.1.11.5).
 GN RCB OR BB0633.
 GN Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Clayton R.A.,
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kertlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Utermack T., Wathley L., McDonald L., Artiach P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi";
 RL Nature 390:580-586(1997).
 CC -1- FUNCTION: REQUIRED FOR EFFICIENT DNA REPAIR: IT CATALYZES THE
 UNWINDING OF DOUBLE-STRANDED DNA AND THE CLEAVAGE OF SINGLE-
 STRANDED DNA AND IT STIMULATES LOCAL GENERIC RECOMBINATION.
 CC ALL OF THESE ACTIVITIES REQUIRE CONCOMITANT HYDROLYSIS OF ATP
 (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
 ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-
 phosphooligonucleotides.
 CC -1- SUBUNIT: CONSIST OF THREE SUBUNITS, RCB, RECC AND RECD
 (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; AE001164; AAC66981.1; -;
 DR HSSP; P56255; 1PJR.
 DR TIGR; BB0633; -;

```

DR InterPro: IPR004586; RecB.
DR InterPro: IPR000212; UvrD-helicase.
DR Pfam: PF00580; UvrD-helicase; 1.
DR TIGERFAMS: TIGR00609; recB; 1.
RW HydroLase; Nuclease; Endonuclease; Helicase; ATP-Binding;
KW DNA repair; Complete proteome.
FT NP_BIND 18 ATP (POTENTIAL).
FT BIND 25
SO SEQUENCE 1169 AA; 137828 MW; B61D63C1C559B91F CRC64;

Query Match      49.0%; Score 49; DB 1; Length 1169;
Best Local Similarity 42.1%; Pred. No. 5.8;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY          1 VERVDFTNHLEDTRRINK 19
           ::::|||||: : : ||
Db          459 IEKIDFTNLSLPQKNNNK 477

RESULT 4
NOCD_HUMAN
ID NOCD_HUMAN STANDARD; PRT; 210 AA.
AC Q43402;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neighbor of COX4.
GN NOCD OR COX4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Retina, and Placenta;
RX MEDLINE=99269920; PubMed=10337626;
RA Bachman N.J., Wu W., Schmidt T.R., Grossman L.I., Lomax M.I.;
RT "The 5-prime region of the COX4 gene contains a novel overlapping
   gene, NOCD."
RL Mamm. Genome 10:506-512(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- TISSUE SPECIFICITY: Expressed in liver, pancreas, heart, lung,
   kidney, brain, skeletal muscle, and placenta. Expression levels
   are highest in pancreas and moderate in heart, skeletal muscle,
   and placenta.
CC -1- SIMILARITY: BELONGS TO THE UPF0172 (NOCD) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
-----
DR EMBL: AF005888; AAB94489.1; -
DR EMBL: AF005889; AAB94489.1; -
DR EMBL: BC005886; AAB05886.1; -
DR EMBL: BC001472; AAB01472.1; -
DR Genew; HGNC:7864; NOCD4.
DR MIM: 604886; -
DR InterPro: IPR005366; UPF0172.
DR Pfam: PF03665; UPF0172.1.
SO SEQUENCE 210 AA; 23773 MW; 41367FDD98769250 CRC64;

Query Match      48.5%; Score 48.5; DB 1; Length 210;
Best Local Similarity 52.4%; Pred. No. 1;
Matches 11; Conservative 2; Mismatches 3; Indels 5; Gaps 1;

OY          4 VDFPNHLEDTRRN----TNK 19

```

```

Db      184 VDFDNHDDIRNDWTNPINRK 204          ||| ||| : | : |||
                                         |||
RESULT 5
NOC4_MOUSE
ID     NOC4_MOUSE        STANDARD:    PRT:    207 AA.
AC     070378;
DT     16-OCT-2001 (Rel. 40, Created)
DT     16-OCT-2001 (Rel. 40, Last sequence update)
DT     15-JUN-2002 (Rel. 41, Last annotation update)
DE     Neighbor of COX4.
GN     NOC4 OR COX4AL.
OS     Mus musculus (Mouse).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX     NCBI_TaxId=10090;
       [1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=C57BL/6J; TISSUE=Kidney;
RC      MEDLINE=99269920; PubMed=10337626;
RX      Bachman N.J., Wu W., Schmidt T.R., Grossman L.I., Lomax M.I.;
RT      "The 5'-prime region of the COX4 gene contains a novel overlapping
RL      gene, NOC4."
RM      Mamm. Genome 10:506-512(1999).
CC      -I- SIMILARITY: BELONGS TO THE UPF0172 (NOC4) FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: AF052621; AAC12933.1; -.
DR      MGD: MGI:1343095; NOC4.
DR      InterPro: IPR005366; UPF0172.
DR      Pfam: PF03665; UPF0172; 1.
SQ      SEQUENCE   207 AA: 23348 MW: 8F3CEP1BF2498958 CRC64;
Query Match              47.5%; Score 47.5; DB 1; Length 207;
Best local Similarity    52.4%; Pred. No. 1.5;
Matches 11; Conservative 2; Mismatches 3; Indels 5; Gaps 1;
OY      4 VDFTNHLEDTRRN-----INK 19
           ||| ||| : | : |||
Db      181 VDFDNHDDIRSDWTNPINRK 201
RESULT 6
CP7A_RABIT
ID     CP7A_RABIT        STANDARD:    PRT:    501 AA.
AC     P51542;
DT     01-OCT-1996 (Rel. 34, Created)
DT     01-OCT-1996 (Rel. 34, Last sequence update)
DT     30-MAY-2000 (Rel. 39, Last annotation update)
DE     Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYPVII)
DE     (EC 1.14.13.17) (Cholesterol 7-alpha-hydroxylase).
GN     CYP7A1 OR CYP7.
OS     Oryctolagus cuniculus (Rabbit).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC     Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX     NCBI_TaxId=9986;
       [1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=New Zealand white; TISSUE=Liver;
RC      MEDLINE=95271192; PubMed=7751825;
RX      Kai M., Eto T., Kondo K., Setoguchi Y., Higashi S., Maeda Y.,
RA      Setoguchi T.;
RT      "Synchronous circadian rhythms of mRNA levels and activities of
RT      cholesterol 7 alpha-hydroxylase in the rabbit and rat.";
RL      J. Lipid Res 36:367-374(1995).
```

```

CC -1- CATALYTIC ACTIVITY: Cholesterol + NADPH + O(2) = 7-alpha-
CC Hydroxycholesterol + NADP(+) + H(2)O.
CC -1- PATHWAY: FIRST AND RATE-LIMITING STEP IN THE CONVERSION OF
CC CHOLESTEROL TO BILE ACIDS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L10754; AAA74382.1; -
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450: 1
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Cholesterol metabolism; NADP.
FT BINDING 441 441 HEME (BY SIMILARITY).
SQ SEQUENCE 501 AA; 58090 MW; FEF1247B151AA6B1 CRC64;

Query Match 47.0%; Score 47; DB 1; Length 501;
Best Local Similarity 66.7%; Pred. No. 4.7;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 5 DFTNHLDDTRRNK 19
Db 366 DFTLHEDGSYNIRK 380

RESULT 7
CP7A_MOUSE STANDARD; PRT; 503 AA.
AC Q64505;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYPVII)
DE (EC 1.14.13.17) (Cholesterol 7-alpha-hydroxylase).
GN CYP7A1 OR CYP7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94375022; PubMed=8086795;
RA Tzung K.W., Ishimura-Oka K., Kihara S., Oka K., Chan L.;
RT "Structure of the mouse cholesterol 7 alpha-hydroxylase gene.";
RL Genomics 21:244-247(1994).
CC -1- CATALYTIC ACTIVITY: Cholesterol + NADPH + O(2) = 7-alpha-
CC hydroxycholesterol + NADP(+) + H(2)O.
CC -1- PATHWAY: FIRST AND RATE-LIMITING STEP IN THE CONVERSION OF
CC CHOLESTEROL TO BILE ACIDS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L23754; AAA68867.1; -
DR MGD: MGI:106091; CYP7A1.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450: 1.
DR PRINTS: PR00385; P450.

```

```

DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Cholesterol metabolism; NADP.
FT BINDING 444 444 HEME (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57258 MW; 51DB3C71D6C79791 CRC64;

Query Match 47.0%; Score 47; DB 1; Length 503;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 5 DFTNHLDDTRRNK 19
Db 369 DFTLHEDGSYNIRK 383

RESULT 8
CP7A_RAT STANDARD; PRT; 503 AA.
AC P18125; P51543;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYPVII)
DE (EC 1.14.13.17) (Cholesterol 7-alpha-hydroxylase).
GN CYP7A1 OR CYP7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Liver;
RC MEDLINE=90307735; PubMed=1694852;
RX Li Y.C., Wang D.P., Chiang J.Y.L.;
RA "Regulation of cholesterol 7 alpha-hydroxylase in the liver. Cloning,
RT sequencing, and regulation of cholesterol 7 alpha-hydroxylase mRNA.";
RL J. Biol. Chem. 265:12012-12019(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RC MEDLINE=90243699; PubMed=2335522;
RX Uehinek D.F., Andersson S., Slaughter C.A., Russell D.W.;
RA "Cloning and regulation of cholesterol 7 alpha-hydroxylase, the rate-
RT limiting enzyme in bile acid biosynthesis.";
RL J. Biol. Chem. 265:8190-8197(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=90033362; PubMed=2806567;
RA Noshiro M., Nishimoto M., Morohashi K., Okuda K.;
RT "Molecular cloning of cDNA for cholesterol 7 alpha-hydroxylase from
RT rat liver microsomes. Nucleotide sequence and expression.";
RL FEBS Lett. 257:97-100(1989).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90277612; PubMed=1693613;
RA Noshiro M., Nishimoto M., Okuda K.;
RT "Rat liver cholesterol 7 alpha-hydroxylase. Pretranslational
RT regulation for circadian rhythm.";
RL J. Biol. Chem. 265:10036-10041(1990).
RN [5]
RP SEQUENCE FROM N.A.
RA MEDLINE=91177904; PubMed=2007596;
RA Nishimoto M., Gotoh O., Okuda K., Noshiro M.;
RT "Structural analysis of the gene encoding rat cholesterol alpha-
RT hydroxylase, the key enzyme for bile acid biosynthesis.";
RL J. Biol. Chem. 266:6467-6471(1991).
RN [6]
RP SEQUENCE OF 1-325 FROM N.A.
RA STRAIN=Sprague-Dawley;
RC MEDLINE=94292507; PubMed=8021257;
RA Chiang J.Y., Stroup D.;
RT "Identification and characterization of a putative bile acid-

```

```

RT responsive element in cholesterol 7 alpha-hydroxylase gene promoter."
RL J. Biol. Chem. 269:17502-17507(1994).
RN [7]
RP SEQUENCE OF 1-27 FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=91084435; PubMed=2261433;
RA Jellinek D.F., Russell D.W.;
RT "Structure of the rat gene encoding cholesterol 7 alpha-hydroxylase."
RL Biochemistry 29:7781-7785(1990).
RN [8]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=93041942; PubMed=1420318;
RA Chiang J.Y., Yang T.P., Mang D.P.;
RT "Cloning and 5'-flanking sequence of a rat cholesterol 7 alpha-hydroxylase gene."
RL Biochim. Biophys. Acta 1132:337-339(1992).
CC -1- CATALYTIC ACTIVITY: Cholesterol + NADPH + O(2) -> 7-alpha-hydroxycholesterol + NADP(+).
CC -1- ENZYME REGULATION: INHIBITION OF ACTIVITY BY BILE ACID FEEDBACK.
CC -1- PATHWAY: FIRST AND RATE-LIMITING STEP IN THE CONVERSION OF CHOLESTEROL TO BILE ACIDS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: Liver.
CC -1- INDUCTION: BY CHOLESTYRAMINE AND CHOLESTEROL.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J05509; AAA40839.1; -
DR EMBL: J02926; AAA40923.1; -
DR EMBL: J05430; AAA41041.1; -
DR EMBL: J05460; AAA03649.1; -
DR EMBL: M59189; AAA41042.1; -
DR EMBL: M59184; AAA41042.1; JOINED.
DR EMBL: M59185; AAA41042.1; JOINED.
DR EMBL: M59186; AAA41042.1; JOINED.
DR EMBL: M59187; AAA41042.1; JOINED.
DR EMBL: M59188; AAA41042.1; JOINED.
DR EMBL: U01962; AAA21144.2; -
DR EMBL: X17993; CAB57878.1; -
DR EMBL: Z14108; CAB78481.1; -
DR PIR: A35376; A35376.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450.1
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme; Microsome; Endoplasmic reticulum; Cholesterol metabolism; NADP.
FT BINDING 444 444 HEME (BY SIMILARITY).
FT CONFLICT 371 371 T -> S (IN REF. 5).
SO SEQUENCE 503 AA; 56882 MW; EA825AA6E74BF5F6 CRC64;

Query Match 47.0%; Score 47; DB 1; Length 503;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

DE Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYP7A1)
DE (EC 1.14.13.17) (Cholesterol 7-alpha-hydroxylase).
DE CYP7A1 OR CYP7.
OS Citellus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; OC Citellus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94029018; PubMed=8105753;
RA Crestani M., Galli G., Chiang J.Y.;
RT "Genomic cloning, sequencing, and analysis of the hamster cholesterol 7 alpha-hydroxylase gene (CYP7)."
RL Arch. Biochem. Biophys. 306:451-460(1993).
CC -1- CATALYTIC ACTIVITY: Cholesterol + NADPH + O(2) -> 7-alpha-hydroxycholesterol + NADP(+).
CC -1- PATHWAY: FIRST AND RATE-LIMITING STEP IN THE CONVERSION OF CHOLESTEROL TO BILE ACIDS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L04690; AAA03751.1; -
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450.1.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme; Microsome; Endoplasmic reticulum; Cholesterol metabolism; NADP.
FT BINDING 444 444 HEME (BY SIMILARITY).
SO SEQUENCE 504 AA; 57447 MW; BDSE521D0C195257 CRC64;

Query Match 47.0%; Score 47; DB 1; Length 504;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 5 DFTNLEDPTRRNINK 19
DB 369 DFTLHLEDGYNIRK 383

CP7A_HUMAN
ID CP7A_HUMAN STANDARD; PRT; 504 AA.
AC P46534;
DC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

CP7A_HUMAN
ID CP7A_HUMAN STANDARD; PRT; 504 AA.
AC P22680;
DC 01-AUG-1991 (Rel. 19, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome P450 7A1 (Cholesterol 7-alpha-monooxygenase) (CYP7A1)
DE (EC 1.14.13.17) (Cholesterol 7-alpha-hydroxylase).
GN CYP7A1 OR CYP7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93176797; PubMed=8439551;
RA Nishimoto M., Noshiro M., Okuda K.;
RT "Structure of the gene encoding human liver cholesterol 7 alpha-hydroxylase."
RL Biochim. Biophys. Acta 1172:147-150(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90346120; PubMed=2384150;

```

RA Noshiro M., Okuda K.;
 RT "Molecular cloning and sequence analysis of cDNA encoding human
 RL cholesterol 7 alpha-hydroxylase.";
 RN FEBS Lett. 268:137-140(1990).
 [3]
 RP SEQUENCE FROM N.A., AND VARIANT SER-100.
 RX MEDLINE-92304280; PubMed-1610352;
 RA Karam W.G., Chiang J.Y.;
 RT "Polymorphisms of human cholesterol 7 alpha-hydroxylase.";
 RL Biochem. Biophys. Res. Commun. 185:588-595(1992).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + NADPH + O(2) = 7-alpha-
 CC hydroxycholesterol + NADP(+) + H(2)O.
 CC -1- PATHWAY: FIRST AND RATE-LIMITING STEP IN THE CONVERSION OF
 CC CHOLESTEROL TO BILE ACIDS.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X56088; CA39568.1; -;
 DR EMBL, M93133; AA58435.1; -;
 DR PIR, JH0659; JH0659.
 DR PIR, S29818; S29818.
 DR GeneW, HGNC:2651; CYP7A1.
 DR MIM: 118455; -;
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Cholesterol metabolism; NADP;
 KW Polymorphism.
 FT BINDING 444 444 HEME (BY SIMILARITY).
 FT VARIANT 100 100 F -> S.
 FT FTID-VAR-001259.
 FT CONFLICT 347 347 D -> N (IN REF. 2).
 FT CONFLICT 385 385 D -> S (IN REF. 2).
 SQ SEQUENCE 504 AA; 57660 MW; D8067E0FF6342949 CRC64;
 Query Match 47.0%; Score 47; DB 1; Length 504;
 Best Local Similarity 66.7%; Pred. No. 4.8;
 Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 5 DFTNHEDETRRNK 19
 DB 369 DFTLHEDGSYNIRK 383
 RESULT 11
 MASP_HUMAN
 ID MASP_HUMAN STANDARD; PRT; 375 AA.
 AC P36952;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Maspilin precursor (Protease inhibitor 5).
 GN SERPINB5 OR P15.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 GN NCBL_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE-94120413; PubMed-8290962;
 RA Zou Z., Antosiewicz A., Hendrix M.J.C., Thor A., Neveu M., Sheng S.,
 RA Rafidi K., Sedor E., Sager R.;
 RT "Maspilin, a serpin with tumor-suppressing activity in human mammary

RT epithelial cells.";
 RL Science 263:526-529(1994).
 RN [2]
 RP SEQUENCE OF 341-360, AND CHARACTERIZATION.
 RX MEDLINE-95318177; PubMed-7797587;
 RA Pemberton P.A., Wong D.T., Gibson H.L., Klefer M.C.,
 RA Fitzpatrick P.A., Sager R., Barr P.J.;
 RT "The tumor suppressor maspin does not undergo the stressed to relaxed
 RT transition or inhibit trypsin-like serine proteases. Evidence that
 RT maspin is not a protease inhibitory serpin.";
 RL J. Biol. Chem. 270:15832-15837(1995).
 CC -1- FUNCTION: TUMOR SUPPRESSOR. IT BLOCKS THE GROWTH, INVASION, AND
 CC METASTATIC PROPERTIES OF MAMMARY TUMORS. AS IT DOES NOT UNDERGO
 CC THE S (STRESSED) TO R (RELAXED) CONFORMATIONAL TRANSITION
 CC CHARACTERISTIC OF ACTIVE SERPINS, IT EXHIBITS NO SERINE PROTEASE
 CC INHIBITORY ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: NORMAL MAMMARY EPITHELIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U04313; AA18957.1; -;
 DR HSSP, P05619; 1HLE.
 DR GeneW, HGNC:8949; SERPINB5.
 DR MIM: 154790; -;
 DR InterPro: IPR000240; Maspilin.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR PRINTS; PR000676; MASPILN.
 DR SMART; SM00093; SERPIN; 1.
 DR PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Glycoprotein; Signal.
 FT CHAIN 1 375
 FT SIGNAL 1 2
 FT ACT_SITE 340 341 REACTIVE BOND (BY SIMILARITY).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 375 AA; 42138 MW; F2D49FCD6AC3DC2 CRC64;
 Query Match 46.0%; Score 46; DB 1; Length 375;
 Best Local Similarity 50.0%; Pred. No. 5;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 OY 1 VERVDFTNHEDETRRNIN 18
 DB 116 LETVDKFKLEETKQGIN 133
 RESULT 12
 MASP_MOUSE
 ID MASP_MOUSE STANDARD; PRT; 375 AA.
 AC F70124;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Maspilin precursor (Protease inhibitor 5).
 GN SERPINB5 OR P15 OR SP17.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 GN NCBL_TaxID-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97206139; PubMed-9132279;

RA Zhang M., Sheng S., Maass N., Sager R.;
 RT "maspin: the mouse homolog of a human tumor suppressor gene inhibits
 RT mammary tumor invasion and motility."; *Mol. Med.* 3:49-59(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strauberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TUMOR SUPPRESSOR. IT BLOCKS THE GROWTH, INVASION, AND
 CC METASTATIC PROPERTIES OF MAMMARY TUMORS. AS IT DOES NOT UNDERGO
 CC THE S (STRESSED) TO R (RELAXED) CONFORMATIONAL TRANSITION
 CC CHARACTERISTIC OF ACTIVE SERPINS, IT EXHIBITS NO SERINE PROTEASE
 CC INHIBITORY ACTIVITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U54705; AAB06042.1; -
 DR EMBL: BC005434; AA05434.1; -
 DR HSSP: P05619; 1HLE.
 DR MGD: MGI:109579; Serpinb5.
 DR InterPro: IPR000240; Maspin.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR PRINTS: PRO0676; MASPIN.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 DR Serpin; Glycoprotein; Signal.
 KW CHAIN 1
 FT SIGNAL 1
 FT ACT_SITE 340 341 REACTIVE BOND (BY SIMILARITY).
 FT CARBOHYD 133 133 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SO SEQUENCE 375 AA; 42111 MW; 5894DDB9A76F9A7 CRC64;
 Query Match 46.0%; Score 46; DB 1; Length 375;
 Best Local Similarity 50.0%; Pred. No. 5;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VERVDFTHNLEDTRRNIN 18
 DB 116 LETVDFKDKLEETKQIN 133
 RESULT 13
 MASP_RAT
 ID MASP_RAT STANDARD: PRT: 375 AA.
 AC -P70564;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Maspin precursor (Protease inhibitor 5).
 GN SERPINB5 OR P15.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Vagina;
 RX MEDLINE=97218145; PubMed=9065806;
 RA Umekita Y., Hilepka R.A., Lleo S.;
 RT "Rat and human maspin: structures, metastatic suppressor activity
 RT and mutation in prostate cancer cells."; *Cancer Lett.* 113:87-93(1997).
 RL

CC -1- FUNCTION: TUMOR SUPPRESSOR. IT BLOCKS THE GROWTH, INVASION, AND
 CC METASTATIC PROPERTIES OF MAMMARY TUMORS. AS IT DOES NOT UNDERGO
 CC THE S (STRESSED) TO R (RELAXED) CONFORMATIONAL TRANSITION
 CC CHARACTERISTIC OF ACTIVE SERPINS, IT EXHIBITS NO SERINE PROTEASE
 CC INHIBITORY ACTIVITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U58657; AAB06043.1; -
 DR HSSP: P05619; 1HLE.
 DR InterPro: IPR000240; Maspin.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; serpin; 1.
 DR PRINTS: PRO0676; MASPIN.
 DR SMART: SM00093; SERPIN; 1.
 DR PROSITE: PS00284; SERPIN; 1.
 DR Serpin; Glycoprotein; Signal.
 KW CHAIN 1
 FT SIGNAL 1
 FT ACT_SITE 340 341 REACTIVE BOND (BY SIMILARITY).
 FT CARBOHYD 99 99 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SO SEQUENCE 375 AA; 42063 MW; E82C4E6A0F22482 CRC64;
 Query Match 46.0%; Score 46; DB 1; Length 375;
 Best Local Similarity 50.0%; Pred. No. 5;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VERVDFTHNLEDTRRNIN 18
 DB 116 LETVDFKDKLEETKQIN 133
 RESULT 14
 SCC2_HUMAN
 ID SCC2_HUMAN STANDARD: PRT: 390 AA.
 AC P48594;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Squamous cell carcinoma antigen 2 (SCCA-2) (Leuplin).
 GN SERPINB4 OR SCCA2.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95241462; PubMed=7724531;
 RA Schneider S.S., Schick C., Fish K.E., Miller E., Pena J.C.,
 RA Treter S.D., Hui S.M., Silverman G.A.;
 RT "A serine proteinase inhibitor locus at 18q21.3 contains a tandem
 RT duplication of the human squamous cell carcinoma antigen gene";
 RT Proc. Natl. Acad. Sci. U.S.A. 92:3147-3151(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96013887; PubMed=7589435;
 RA Barnes R.C., Worrall D.M.;
 RT "Identification of a novel human serpin gene: cloning sequencing and
 RT expression of leuplin."; *FEBS Lett.* 373:61-65(1995).
 RL [3]


```

RP SEQUENCE FROM N.A.
RX MEDLINE-21167379; PubMed-11267667;
RA Hamada K., Shinomiya H., Asano Y., Kihana T., Iwamoto M., Hanakawa Y.,
RA Hashimoto K., Hirose S., Kyo S., Ito M.;
RT "Molecular cloning of human squamous cell carcinoma antigen 1 gene and
RT characterization of its promoter."
RL Biochim. Biophys. Acta 1518:124-131(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY ACT AS A PROTEASE INHIBITOR TO MODULATE THE HOST
CC IMMUNE RESPONSE AGAINST TUMOR CELLS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: SQUAMOUS CELLS.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19576; AAA92602.1; -.
DR EMBL: U19570; AAA92602.1; JOINED.
DR EMBL: U19571; AAA92602.1; JOINED.
DR EMBL: U19572; AAA92602.1; JOINED.
DR EMBL: U19574; AAA92602.1; JOINED.
DR EMBL: U19575; AAA92602.1; JOINED.
DR EMBL: U19557; AAA97553.1; -.
DR EMBL: X89015; CAA61420.1; -.
DR EMBL: AB035089; BAB21525.1; -.
DR EMBL: BC017401; AAH17401.1; -.
DR HSSP: P01008; IATH.
DR Genew: HGNC:10570; SERPINB4.
DR MIM: 600518; -.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF000079; serpin; 1.
DR SMART: SM00093; SERPIN; 1.
DR PROSITE: PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT_SITE 354 355 REACTIVE BOND.
SQ SEQUENCE 390 AA; 44854 MW; 0AE213CD892587D5 CRC64;

Query Match 46.0%; Score 46; DB 1; Length 390;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 VERVDFTNHLEDTNRNIN 18
Db 131 VESTDFRANPEESRKIN 148

RESULT 15
BXG_CLOBO STANDARD; PRT; 1296 AA.
ID 060393;
AC 060393;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
DE (Bontoxilysin G).
GN BONTG.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;

```

```

RX MEDLINE-94092745; PubMed-8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
RT (Clostridium argentineense) type G neurotoxin: genealogical comparison
RT with other Clostridial neurotoxins."
RL Biochim. Biophys. Acta 1216:487-491(1993).
RN [1]
RP SEQUENCE: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexcitatory apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74162; CAA52275.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXISIN.
DR Prodom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 1 441
FT CHAIN 0
FT CHAIN 1 441
FT CHAIN 442 1296
FT METAL 229 229
FT METAL 230 230
FT ACT_SITE 233 233
FT METAL 435 449
FT DISULFID 435 449
SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F65C31 CRC64;

Query Match 46.0%; Score 46; DB 1; Length 1296;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VERVDFTNHLEDTNRNIN 19
Db 554 IENILQTLNSLDALRNK 572

```

Search completed: July 11, 2003, 11:55:14
Job time : 3.81705 secs